POSTER

2019

PROCEEDINGS OF THE INTERNATIONAL STUDENT SCIENTIFIC CONFERENCE POSTER – 23/2019

PROCEEDINGS OF THE INTERNATIONAL STUDENT SCIENTIFIC CONFERENCE POSTER – 23/2019

The objective of the international scientific student conference POSTER 2019 is to provide an international forum for presentation of undergraduate and postgraduate student research work in fields related to electrical engineering.

All contributions were presented as posters in eight parallel sections at the Faculty of Electrical Engineering, Czech Technical University in Prague on May 23, 2019.

ISBN 978-80-01-06581-5

Editor: Libor Husník

Proceedings of the International Student Scientific Conference Poster – 23/2019

Publisher: Czech Technical University in Prague

Processed by: Faculty of Electrical Engineering, CTU in Prague Technická 2, 166 27, Prague 6, Czech Republic, tel: +420 22435 2115

Printed by: Faculty of Electrical Engineering, Printing Office, CTU in Prague

Nr. of pages: 86

Issue: 1

The international student conference POSTER 2019, organized by the Faculty of Electrical Engineering, Czech Technical University in Prague, started as an internal meeting of doctoral students in 1995. In 1997, the first International Student Conference POSTER was held. On May 23, 2019 we organized the Conference for the 23rd time, this year in cooperation with the Faculty of Biomedical Engineering and Faculty of Information Technology.

One of the main goals of the POSTER conference is to promote discussions and interactions among postgraduate and undergraduate students from various countries, universities and fields of study related to electrical engineering. Another important goal is to support independent creative work of students and stimulate practical application of acquired theoretical knowledge since we consider the students' research activity an inevitable part of the whole educational process at all universities.

This year we introduced the following change to the conference. Students can submit either a two-page abstract or a four-page full paper for publication. They can also opt out of publication of their contributions.

The program committee of Poster 2019 conference selected 95 papers for presentation at the conference from the total of 97 submitted contributions. The contributions are divided into eight sections:

- Biomedical Engineering
- Communications
- Electronics and Instrumentation
- History of Science
- Informatics and Cybernetics
- Management
- Natural Sciences
- Power Engineering

A total of 60 contributions come from Czech Technical University in Prague, 2 from other Czech universities and 33 contributions are from abroad. Criteria of acceptance were based on the originality of scientific contribution and good English. Each paper was reviewed by at least two members of the respective section of the Programme Committee. A total of 9 authors opted out of the publication of their contributions.

We would like to thank all students who produced outstanding research results and contributed to this proceeding.

Last but not least, we would like to express our gratitude to all colleagues from the Office for Science, Research and International Relations and from the Computing and Information Centre who helped us a lot in preparation of the proceedings and organization of the conference.

Prague May 2, 2019

Libor Husník
on behalf of the Organizing Committee of POSTER 2019

List of papers

Paper ID	Paper Title		First Author	Other Authors
BI01	Medical Applications of System Identification with Pseudo Random Binary Sequences	Abstract	Stephan Eisenbrand	Tobias Menden
BI02	3D simulation to analyse the effect of thrombosis on bioimpedance spectroscopy	Abstract	Stephan Dahlmanns	Leon Geller-Hartung
BI03	Testbench to model cardiac volume changes	Abstract	Leonie Korn	Max Rumpf
BI04	Organic LED panels for pulse rate measurement using photoplethysmography imaging	Full paper	Simon Lyra	Michael Paul
BI05	High frequency oscillation analysis in epileptic patients	Abstract	Martin Dostál	
BI06	Remote breathing signal extraction using infrared thermography in a climatic chamber	Abstract	Christoph Weiss	Alexander Kirmas
BI07	Feedforward Compensation of Physical Activity in Insulin Therapy for Type1 Diabetes using Oxygen Uptake	Full paper	Bianca Romanski	Carl-Friedrich Benner
BI08	Glycemic Control in the Intensive Care Unit to prevent Hyperglycemia in Critically Ill Patients	Abstract	Carl-Friedrich Benner	
BI09	Towards Subjective Evaluation of Individualized Head-Related Transfer Functions	Abstract	Shaima'a Doma	Hark Braren
BI10	Subject Movement During the Measurement of Head-Related Transfer Functions	Abstract	Saskia Wepner	Jan-Gerrit Richter
BI11	Signal recording of 3D neurospheres on high-resolution CMOS MEA platform	Abstract	Steffen Künzinger	
BI12	Functionalized Hydrogel as Glucose Imaging Sensor for in vitro Cell Cultures	Abstract	Jens Jüttner	Lukas Kress
BI13	ICA decomposition as a basis for finding hidden EEG structures by dimension reduction methods	Full paper	Natálie Brožová	
BI14	Airwave Oscillometry	Abstract	Tomáš Vlček	
BI15	Assessment of Tidal Volume in High-frequency Oscillatory Ventilation		Marek Kordík	
C01	Analytical Review of Erbium-Ytterbium Waveguide Amplifier for C-Band Allocated DWDM System	Abstract	Tomáš Huszaník	
C02	Analysis of Reflector Size Impact on Modeled Helix Antenna for 5.2 GHz Usable for Experimental FSO/RF Hybrid System	Abstract	Michal Márton	Michal Špes, Jakub Urbanský
C03	Binarization of Feature Vectors for Image Retrieval	Abstract	Erik Ottlik	Abin Jose
C04	Hierarchical Depth Map Estimation for Inter-Prediction in Video Coding	Full paper	Ishan Saxena	Hossein Bakhshi-Golestani
C05	Measurement of Rotational Degrees of Freedom in Vibroacoustic Transfer Functions		Christian Dreier	Mark Müller-Giebeler
C06	Simulation and fabrication of 3D printed optical planar multimode 1x2Y splitters		Jakub Hybler	Jakub Dlouhy
EI01	Modeling of a Hump Effect Using a Three-Dimensional TCAD Device Simulator	Abstract	Dalibor Barri	Dalibor Barri
EI02	A procedure for frequency selection of multisine impedance spectroscopy regarding higher harmonics at undersampling conditions	Abstract	Alexander Blömeke	
EI03	Selectivity enhancement of target volatile organic compounds using sensor array and temperature modulation	Full paper	Ali Ahmad	
EI04	High Performance Ultra-low Quiescent Current LDO Regulator	Full paper	Lukas Buryanec	
EI05	Design and Analysis of an Interdigital Capacitor in Microstrip Line Technology for Low Noise X-band System Applications	Abstract	Kai Fan	
EI06	Parameterisation of an Equivalent Circuit Model of an LTO Battery Cell		Maya Strautmann	Gereon Stahl
EI07	Audio Quality Assessment of Digital Radio Broadcasting Systems	Abstract	Adam Bartyzal	
EI08	All-Inkjet Printed Flexible Capacitors Utilizing SU-8 and PVP Dielectric Inks	Abstract	Vojtěch Povolný	
EI09	Radiation Tolerant 8-bit Analog to Digital Converter with Successive approximation	Abstract	Pavel Vancura	
EI10	Characterizaton of LVDS drivers for MAPS detector designed in 180 nm SOI CMOS technology	Abstract	Tomas Benka	
EI11	Numerical optimization of coils system geometry for homogeneous magnetic field generation	Full paper	David Novotný	
HS01	Introduction of Photography Teaching at the Imperial and Royal Czech Technical University in Prague	Abstract	Pavel Borkovec	
HS02	Thundering the Engines of Flying Fortresses and Liberators over the “Palaba” Factory	Full paper	Zbyněk Nikel	
HS03	The Harvester SM-500 – the machine, which brought joy and sorrow	Full paper	Martin Vychytil	
HS04	History of shoe-making in Skuteč up to 1949	Full paper	Frantisek Krejci	

Paper ID	Paper Title		First Author	Other Authors
HS05	Amazonia and the hydropower plants: urban networks, decrease of the forest and spatial inequalities	Abstract	Nádia Matioli Yazbek Bitar	
HS06	Gaius Sergius Orata and the Underfloor Heating System		Martin Pola	
HS07	Jan Henniger of Eberk, Lawyer at the Services of the Prague Polytechnic	Abstract	Vít Šmerha	
HS08	The Masaryk Academy of Labour and its Contribution to the Development of Technically Skilled Workers	Abstract	Vít Holeček	
HS09	History of Teletext Services in Context of Hearing Impaired People	Abstract	Zdenek Bumbalek	
HS10	The development of DNA microarrays	Abstract	Vladimír Kunc	
HS11	Personal sources relating to the electrotechnical education at VŠB	Full paper	Mariana Stonišová	
HS12	History of transmissions in Mlada Boleslac	Full paper	Lukáš Nachtmann	
IC01	Looking for existing relations and semantics between test cases and production source code	Full paper	Matej Madeja	
IC02	Approximate search in genomic data	Full paper	Petr Ryšavý	
IC03	Hybrid Code Networks using a convolutional neural network as an input layer achieves higher turn accuracy	Full paper	Petr Marek	
IC04	A Simulink-Blockset to Control Lego Mindstorms using Raspberry Pi	Abstract	Isshu Araki	Benjamin Hentze
IC05	Datasets Generation for Neural-Symbolic Integration	Abstract	Martin Svatoš	
IC06	System for automatic music genre recognition	Abstract	Maroš Jakubec	
IC07	A simple acoustic surveillance system for child's emotion detection	Abstract	Eva Lieskovská	
IC08	On the importance of dropout for Checkerboard D-GEX architectures	Abstract	Vladimír Kunc	
IC09	Design of semantic model for data mining processes	Full paper	Juliana Ivančáková	Michal Ivančák
IC10	Web-based collaborative virtual environments to support cross-platform access	Full paper	Marián Hudák	Martin Sivý
IC11	Entity Recognition Using Contextual Embeddings	Abstract	Jan Pichl	
IC12	Eliminating Semantic Ambiguities Using Intensional Logic	Abstract	Zuzana Bilanová	Matúš Uchnár, Martin Štancel
IC13	Uniformed Intelligent user interface for Smart Home using Virtual Reality	Full paper	Martin Sivý	Marián Hudák
IC14	CNN based near-duplicate image detectionused for email spam detection	Full paper	Vít Listík	
IC15	VeraGreg: A Framework for Verifiable Privacy-Preserving Data Aggregation		Jakub Klemsa	Jan Říha
IC16	Using sentiment detection on tweets to form opinions	Abstract	Jakub konrád	
IC17	Automatic Test of Visual Feature Detectors and Descriptors for Visual Odometry	Full paper	Tomáš Pivoňka	
IC18	Optimal evaluation of risky opportunities	Abstract	Matej Uhrín	
M01	Weighting Indicators in Composite Indexes: a Theoretical Model of Competitiveness Assessment for Electronics Market	Full paper	Mindaugas Samoška	Mindaugas Samoška
M02	Economic Evaluation of FDM 3D Printing Method	Abstract	Daniel Běhal	Tomáš Tichý, Lukáš Dvořáček
M03	LPS For Power Supply facility	Abstract	Vladimir Konusarov	
M04	Decentralized electricity supplies of remote area	Full paper	Vladislav Shlapak	
M05	Development of a hybrid power plant for the power supply of a rural area	Full paper	Alena Pavlova	
M06	Hybrid Power Supply in the Koyda Village	Full paper	Andrey Kurganovskiy	
M07	Evaluation of the effectiveness of photovoltaic modules	Full paper	Mikhail Ignatenko	
M08	Benefits of Battery Storage for Charging Stations	Full paper	Lukáš Dvořáček	
M09	Prediction of household consumption based on the model	Full paper	Martin Horák	
M10	Comparative analysis of micro cogeneration for family house in the Czech Republic and the Russian Federation	Full paper	Viktoriia Pereboeva	
M11	Finantial Time Series Forecasting With Wavelets	Abstract	Jaroslav Schürrer	
NS01	Polyaniline electric conductivity in the presence of ammonia molecules	Full paper	Hana Šustková	
NS02	Characterization of plasma in the PF-1000 plasma focus experiments	Full paper	Pavla Bednářová	Lucia Čabrová
NS03	Optimizing of experimental load of PFZ-200 plasma focus	Full paper	Jan Novotný	
NS04	Experimental System for Classification of Defects in Vinyl Disc Records based on Typical Defects Attributes: First Proposal	Full paper	Václav Moldan	

Paper ID	Paper Title		First Author	Other Authors
NS05	In-situ Raman spectroscopy characterization of SU-8 epoxy resin temperature dependent curing process	Abstract	Alexandr Posta	
NS06	Virtual Acoustic Space Test with HMD: Pilot Test Proposal Using Unity to Create Suitable Test Structure	Full paper	Viktor Jarolímek	
NS07	Measuring otoacoustic emissions		David Šlegl	Ondřej Klimeš
PE01	Observer-Based Identification of Nonlinear Magnetizing Inductance of Induction Machine	Abstract	Ondrej Lipcak	
PE02	Improvement of transient stability using STATCOM with a nonlinear PID controller	Full paper	Lavr Vetoshkin	
PE03	A new approach to investigation of Series Arc behaviour	Full paper	Lavr Vetoshkin	
PE04	Effective calculation of heat flow trough walls	Full paper	Jan Rimbala	
PE05	Model of operation modes of a stand-alone photovoltaic systems taking into account the actual operating conditions	Full paper	Timofei Abramov	
PE06	Evaluation of substrate cleaning process based on contact angle measurement	Abstract	Jakub Jirsa	
PE07	Implementation of Small Wind Turbine in Matlab Simulink	Full paper	Jakub Urbanský	Michal Špes, Daniel Pál, Michal Márton
PE08	Analysis of integration of accumulation systems in Smart Grid	Full paper	Maksym Oliinyk	Daniel Pál
PE09	Model of the microgrid for the future smart grid	Full paper	Michal Ivančák	Juliana Ivančáková
PE10	Simulation of daily, artificial and combined lighting in the software DIALux	Full paper	Daniel Pál	Jakub Urbanský, Maksym Oliinyk
PE11	Design of a measurement assembly and methodology for determining the conductor's current load	Full paper	Michal Špes	Jakub Urbansky, Daniel Pál, Michal Márton
PE12	Determination of MMC submodule capacitor value		Jakub Zedník	Tomáš Košťál
PE13	Synthesis and Spark Plasma Sintering of CaCu3Ti4O12	Abstract	Jan Petrášek	
PE14	An introduction to voltage gain of LLC resonant converter	Abstract	Filipp Frolov	
PE15	Surface Defect Spectroscopy of Transparent Conductive Oxides		Eva Horynova	

Medical Applications of System Identification with Pseudo Random Binary Sequences

Stephan EISENBRAND[1], Tobias MENDEN[1]

[1]Medical Information Technology, RWTH Aachen University, Pauwelsst. 20, 52074 Aachen, Germany

stephan.eisenbrand@rwth-aachen.de, menden@hia.rwth-aachen.de

Abstract. *For real time system identification random and pseudo-random signals can be used to greatly reduce measurement time and to grasp physiological parameters faster than using the common method of sinusoidal input signals or frequency sweeps. The most basic random signal is white noise, which leads to a perfect reconstruction of the system response for an infinite measurement time. The practical usage of pseudo-white signals stimulate a physiological system at multiple frequencies at once and reduce frequency leakage errors at the same time. This review explains the basic idea of pseudo-random binary sequence (PRBS) stimulation and shows several practical applications.*

Keywords

pseudo-random binary sequence, PRBS, white noise stimulation, system identification, correlation response, impulse response.

1. Introduction

Physiological systems are in most cases of non-linear nature and have time-variant parameters. Traditionally, sinusoidal signals are used to stimulate the system and tissue parameters are derived from the measured impedance. Frequency sweeps with sinusoidal stimuli take at least the periodic time of each frequency [1], so the time-variant parameters might have changed during the measurement. However, a rarely used approach is the use of noise stimuli to obtain the frequency response of a system. The idea behind the White Noise Approach is to face non-linearity by testing a system with a great variety of stimuli and record how the system responds to them. If a system is excited by white noise for a long enough time, there is a finite probability that any given stimulus is covered by a part of the white noise signal. This fact leads to the equivalence of two systems, if and only if their system response is the same to white noise stimulation. Accordingly, an adequate mathematical description, which responds in the same way to the noise stimulus as its physiological system can be found [2]. The impedance of the system can be obtained by subsequent application of Fourier Transform. Gaussian White Noise (GWN) theoretically contains every possible frequency and all amplitudes, which speeds up the measurement by testing all frequencies at the same time [3]. In addition the white noise excitation approach is only sparsely affected by extraneous contaminating noise in physiological systems [4]. The practical drawback of this method is the existence of leakage errors, due to the finite measurement time (only infinite measurement time could cover all frequencies) [5]. Therefore, the usage of pseudo-random binary sequences (PRBS) provides a reasonable alternative. PRBS are periodic, but have nearly the same auto-correlation function as random binary sequences and nearly the largest amplitude density. Therefore, PRBS are well suited for the medical field, since low signal to noise ratios and short stable lifetimes are predominant in physiological experiments [2]. Myocardial regeneration, cell characterization in micro-fluidic systems and dynamic electrical impedance tomography are examples of applications, where physiological parameters have to be obtained in a short time and where a standard frequency sweep would not be sufficient to grasp the crucial parameters. However, these approaches have lost significance in the 70's because of the major drawback of insufficient memory and computing speed. Nowadays this approach could present a useful alternative to the traditional system identification through sinusoidal stimuli.

2. Basic Principles

2.1. Correlation and Impulse Response

One approach to obtain the frequency profile of a system is the cross-correlation method. The correlation technique is especially useful for non-repeatable signals or time-varying parameters. Also, extraneous noise, which is always present in real measurements, can be reduced by the correlation method by averaging the signal over a period of time. At first the correlation function is estimated and then its Fourier Transform is computed to get the system response. The basic equations are given by

$$R_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t-\tau)x(t)dt \qquad (1)$$

and

$$R_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t-\tau)y(t)dt, \tag{2}$$

where $R_{xx}(\tau)$ denotes the auto-correlation function and $R_{xy}(\tau)$ the cross-correlation function of an input signal $x(t)$ and an output signal $y(t)$. Furthermore, the cross-correlation function is related to the autocorrelation function by the following convolution integral

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} g(\tau-\lambda)R_{xx}(\lambda)d\lambda, \tag{3}$$

where $g(\tau)$ is the weighting function of the related system. In the frequency domain the power-density function $S_{xx}(j\omega)$ and the cross-spectral density function $S_{xy}(j\omega)$ can be obtained through Fourier Transformation of R_{xx} and R_{xy}:

$$S_{xx}(j\omega) = \int_{-\infty}^{\infty} R_{xx}(\tau)e^{j\omega\tau}d\tau, \tag{4}$$

$$S_{xy}(j\omega) = \int_{-\infty}^{\infty} R_{xy}(\tau)e^{j\omega\tau}d\tau = G(j\omega) * S_{xx}(j\omega). \tag{5}$$

$S_{xx}(j\omega)$ becomes constant, when applying white noise at the input $x(t)$ [4]. Therefore, the cross-spectral density function is directly proportional to the the transfer function $G(j\omega)$ of the system, which implies a low time-consuming caption of $G(j\omega)$ by stimulating the system with white noise.

2.2. Pseudorandom Binary Sequence (PRBS)

Perfect white noise is impossible to generate, due to the infinite bandwidth. However, pseudo-random binary sequences (PRBS) are deterministic binary signals with similar properties to a white noise signal, which are reproducible and therefore well-suited for measurements. PRBS signals are most commonly used for broadband system identification. The quasi-white noise auto-correlation properties remain, whereas the natural redundancy (repetition of waveform-sections in a signal) of the random quasi-white noise process is reduced. This increases the efficiency of the test-signal, since more different signals can be applied to the system in the same time [2].

To compare different kinds of signals, the crest factor (CF) is introduced here. It describes how much signal amplitude is needed to put a certain power into a system. The formula for the CF of a given input signal $x(t)$, exciting a system is given by the ratio between the peak value of input signal x_{peak} and its effective root mean square value x_{rms}:

$$\mathrm{CF}(x) = \frac{x_{\mathrm{peak}}}{x_{\mathrm{rms}}}. \tag{6}$$

Ideally the crest factor is as low as possible and reaching its theoretical lower bound at 1. According to *Allemang et al.* (cf. [6]), eq. 6 can be transformed into the following relation 7 by using Parseval's Theorem. It also includes the fact that all excitation power, even outside the measurement bandwidth, contribute to the excitation of the system, but will not help to get more information out of the system:

$$\mathrm{CF}(x(t)) = \frac{\max_{t\in[0,T]} |x(t)|}{\sqrt{\frac{1}{T}\int_0^T x^2(t)dt} \cdot \sqrt{\frac{P_{\mathrm{int}}}{P_{\mathrm{tot}}}}}. \tag{7}$$

Binary sequences have a far lower crest factor than Gaussian White Noise. Nevertheless, with random binary noise, no proper control over the spectrum is achieved, since the signal is generated out of the filtered sign of a GWN signal. This proceeding distorts the spectrum of the input sequence [7]. Hence, pseudo-random binary sequence signals are used, which are generated differently using a linear-feedback shift register (LFSR). The maximum length of an n-bit PRBS shift register is $2^n - 1$ sequences long. In theory, these sequences contain all frequencies and have the lowest crest factor, compared to other random excitation signals. Compared to a random binary signal, the frequency spectrum of a PRBS can be shifted by changing the clock sampling rate of the LFSR. The drawback of the PRBS signal is the poor identification of non-linear systems, whereas it only consists of two levels, which might not provide enough information to identify non-linearities in a system. On the other hand, a PRBS signal with a very high input maximum will drive the measurement of a linear system into non-linear behavior. Therefore, for non-linear systems, variants of the PRBS are used. Multi-Valued PRBS, Amplitude-Modulated PRBS and Multi-Level PRBS (m-level PRS) [8, 9] can be chosen to have more levels in the input signal and consequently more information about the system [7].

The PRBS signal can be calculated by using the modulo 2 operator to calculate the binary remainder of the weighted shift register signal

$$x[k] = \mathrm{mod2}(a_1x[k-1] + ... + a_n[k-n]), \tag{8}$$

where n is called the cycle length. The clock or cycle time of the shift register is denoted by λ and for small λ the impulse response of the system with input $x(t)$ and output $y(t)$ can be calculated by

$$g(t) = \frac{1}{a^2\lambda^2}R_{xy}(t), \tag{9}$$

where a is the amplitude of the PRBS signal forming the power spectral density over a sufficiently large frequency spectrum [10] by

$$S_{x0} \approx a^2\lambda. \tag{10}$$

There are three free parameters a, λ and n of the PRBS signal, which can be adjusted to modify the input signal of the system. To avoid disturbance in the output signal, the amplitude should be chosen as large as possible, taking into account the input and output range of the system. Also the cycle time λ should be selected as large as possible to get a high spectral density. However, the cutoff frequency ω_c of the PRBS signal is given by

$$\omega_c = \frac{1}{\lambda}, \tag{11}$$

so λ should not be chosen too large and a value of

$$\lambda \leq \frac{T_i}{5}, \tag{12}$$

is suggested [10], where T_i denotes the smallest time constant in the system under test. The length of the cycle n can be set under the assumption that the period time $T_\mathrm{p} = n\lambda$ should not be smaller than the transient time T_{95} of the system, so that no overlap of the impulse responses can occur. A rule of thumb is given by

$$T_\mathrm{p} \approx 1.5 T_{95} \tag{13}$$

3. Applications of PRBS

3.1. Measurement of Blood Glucose

A Recent study from *Guarin et.al.* (2013) [11] presented a concept for measuring glucose concentration in an aqueous solution. The non-invasive method aims for measuring blood glucose levels, which might be very useful for diabetes mellitus patients. They used a micro-strip sensor to measure the glucose concentration. The electromagnetic field is not only present in the micro-strip waveguide, but has also a fringe field, which can interact i.e. with a tissue, when applied to it. Fig. 1 demonstrates the electromagnetic

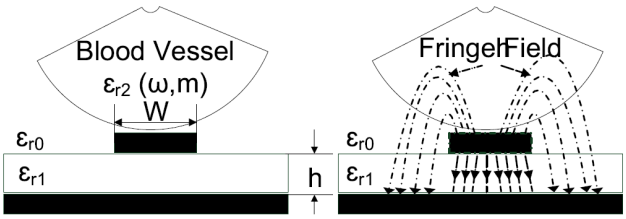


Fig. 1. Microstrip fringefield interacts with a blood vessel [11].

synergy with blood, when the micro-strip sensor is put onto the skin. The interaction can be modeled by the effective permittivity of the waveguide [12]

$$\varepsilon_\mathrm{eff} = \varepsilon_{r1} q_1 + \varepsilon_{r2} \frac{(1-q_1)^2}{\varepsilon_{r2}(1-q_1-q_2)+q_2}, \tag{14}$$

where ε_{r1} depicts the permittivity of the micro-strip and ε_{r2} the permittivity of blood. q_1 and q_2 are filling factors of the micro-strip line. Furthermore, ε_{r2} can be described as a Debye model [13]

$$\varepsilon_{r2} = \varepsilon''_{r2}(\omega, m) + j\varepsilon'_{r2}(\omega, m), \tag{15}$$

since blood can be seen as an aqueous solution. Formula 15 connects the glucose concentration m with frequency ω. ε''_{r2} denotes the real part and ε'_{r2} is the imaginary part of ε_{r2}. Furthermore the attenuation-constant and the wave-number are also related parameters to the level of glucose. The system is assumed to be linear. Stimulating it with a PRBS signal, the wave-related parameters can all be measured by cross-correlating the input signal with the measured output signal, which leads to the impulse response of the system in time domain. Cross-correlation was performed using Hadamard Transformation, which is suitable for correlating Reed-Muller codes, hence applicable to PRBS signals. Using Fast Fourier Transformation the transfer function can be obtained in the frequency spectrum and therefore the blood glucose concentration can be related through frequency and magnitude of the transfer function.

3.2. Impedance-Based Microflow-Cytometry

PRBS signals are sensitive to time variances in a system, which leads to a reduction of the SNR. *Sun et al.* [14] have used an adaptive line enhancer (ALE) in impedance-based micro-flow cytometry for single cell analysis for increasing the SNR again, while stimulating the system with PRBS-signals. The system is based on micro-channels with two measuring chambers with integrated electrodes. One channel serves as reference volume, whereas the cells under test are flowing through the other chamber. The electrical impedance of each single cell can be measured as it flows between the electrode-pairs. The differential signal is composed out of one positive and one negative peak, when the cell flowing through changes the properties of the electromagnetic measurement field in the test chamber. The impedance parameters can be estimated from the peaks of this bipolar signal. The measurement setup in micro-fluidic channels is very dynamic, as the system information of the cells have to be extracted within milliseconds at multiple frequencies while the cells are flowing through the test chamber. This makes it a good example for using broadband random signals.

3.3. Bio Impedance Spectroscopy

Another PRBS-related excitation method was found by *Min et al.* [15], while comparing several broadband stimuli, especially with regard to Bio Impedance Spectroscopy (BIS) applications. They evaluated rectangular pulses, sinc and chirp signals. The sine wave-based signum chirp test-signal showed good result for a well-defined frequency range, crest factor and signal to noise ratio. The general idea from waveform generation to impedance spectrum is depicted in Fig. 2. $\dot{Z}$ is the system under test and comprises the impedance of

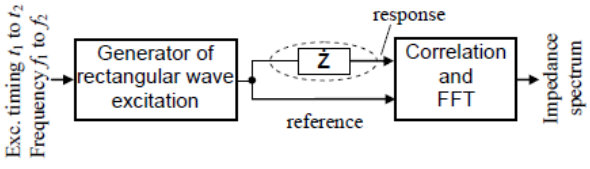


Fig. 2. Structure of the measurement system for rectangular wave chirp [15].

the biological tissue, which is excited by the signal. The system response is cross-correlated in time-domain with the reference-signal and finally mapped into frequency domain via Fourier Transformation. The chirp is a non-periodic one-shot signal with continuously changing frequency, which is generated during a time of $T_{\mathrm{exc}} = 10\,\mathrm{ms}$ and covers a frequency spectrum of approximately $10^5\,\mathrm{Hz}$. The rectangular chirp has an even lower crest factor than the PRBS signal approach generated by an LFSR and is therefore even more effective to use. This approach has been developed in particular for medical applications.

4. Conclusion

In this review, the principle of system analysis with pseudo-random noise signals as system excitation have been presented. The advantage over commonly used sinusoidal stimulus signals is the excitation of many frequencies at once in a much shorter time. White noise would be the best test signal, since it theoretically contains all excitable frequencies. In practice, PRBS-signals are most commonly used, because of their continuous spectrum and low CF. They can be generated very easily by the use of a feed back shift register and have white-noise-like properties. System identification with PRBS input signals are ideally suited for linear systems, but have major drawbacks when it comes to non-linear systems, because they only consist of 2 amplitude levels. Alternative PRBS signals like multi-valued PRBS, Amplitude-Modulated PRBS and multi-level PRBS (m-level PRBS) can be used to cope with non-linear system identification. PRBS has a finite bandwidth and can be easily set to a certain frequency range by changing the clock signal λ of the shift register.

Acknowledgements

This review was supervised by Univ-Prof. Dr.-Ing. Dr. med. Dr. h. c. Steffen Leonhardt and Dr. Marian Walter, RWTH Aachen University.

References

[1] SCHOUKENS, J., PINTELON, R., ROLAIN, Y., Broadband versus stepped sine frf measurements, *IMTC/99. Proceedings of the 16th IEEE Instrumentation and Measurement Technology Conference (Cat. No.99CH36309)*, vol. 2, 1999 pp. 1070–1074 vol.2.

[2] MARMARELIS, P., MARMARELIS, V., *Analysis of Physiological Systems: The White-noise Approach*, Computers in biology and medicine, Kluwer Academic Pub, 1978.

[3] GABRIELLI, C., HUET, F., KEDDAM, M., LIZEE, J., Measurement time versus accuracy trade-off analyzed for electrochemical impedance measurements by means of sine, white noise and step signals, *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry*, 1982, vol. 138, no. 1, pp. 201 – 208.

[4] ICHISE, M., NAGAYANAGI, Y., KOJIMA, T., Application of pseudo-random signals and cross-correlation techniques in electroanalytical chemistry, *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry*, 1974, vol. 49, no. 2, pp. 187 – 198.

[5] SCHOUKENS, J., ROLAIN, Y., PINTELON, R., Leakage reduction in frequency-response function measurements, *IEEE Transactions on Instrumentation and Measurement*, 2006, vol. 55, no. 6, pp. 2286–2291.

[6] ALLEMANG, R., DE CLERCK, J., NIEZRECKI, C., BLOUGH, J., *Topics in Modal Analysis I, Volume 5: Proceedings of the 30th IMAC, A Conference on Structural Dynamics, 2012*, Conference Proceedings of the Society for Experimental Mechanics Series, Springer New York, 2012.

[7] TANGIRALA, A.K., System identification, 2013, lecture, Department of Chemical Engineering IIT Madras.

[8] BRAUN, M.W., RIVERA, D.E., STENMAN, A., FOSLIEN, W., HRENYA, C., Multi-level pseudo-random signal design and ldquo;model-on-demand rdquo; estimation applied to nonlinear identification of a rtp wafer reactor, *Proceedings of the 1999 American Control Conference (Cat. No. 99CH36251)*, vol. 3, 1999 pp. 1573–1577 vol.3.

[9] GODFREY, K. (editor), *Perturbation Signals for System Identification*, Prentice Hall International (UK) Ltd., Hertfordshire, UK, UK, 1993.

[10] ISERMANN, R., MÜNCHHOF, M., *Identification of Dynamic Systems: An Introduction with Applications*, Advanced Textbooks in Control and Signal Processing Series, Springer Berlin Heidelberg, 2010.

[11] GUARIN, G., HOFMANN, M., WEIGEL, R., FISCHER, G., KISSINGER, D., Determination of sugar concentration in aqueous solutions using ultra-wideband microwave impedance spectroscopy, *2013 IEEE MTT-S International Microwave Symposium Digest (MTT)*, 2013 pp. 1–4.

[12] SVACINA, J., Analysis of multilayer microstrip lines by a conformal mapping method, *IEEE Transactions on Microwave Theory and Techniques*, 1992, vol. 40, no. 4, pp. 769–772.

[13] HOFMANN, M., BLOSS, M., WEIGEL, R., FISCHER, G., KISSINGER, D., Non-invasive glucose monitoring using open electromagnetic waveguides, *2012 42nd European Microwave Conference*, 2012 pp. 546–549.

[14] SUN, T., GREEN, N.G., MORGAN, H., Adaptive line enhancer assisted single cell identification in a pseudorandom noise-stimulated microflow-cytometry, *2008 IEEE Sensors*, 2008 pp. 630–633.

[15] MIN, M., PAAVLE, T., ANNUS, P., LAND, R., Rectangular wave excitation in wideband bioimpedance spectroscopy, *2009 IEEE International Workshop on Medical Measurements and Applications*, 2009 pp. 268–271.

About Authors...

Stephan EISENBRAND was born in Dernbach, Germany and received the B.Sc. degrees from RWTH Aachen University, Aachen, Germany in 2017. His research interests include measurement and control topics in medicine.

Tobias MENDEN was born in Bad Honnef, Germany. 2016 he received the M.Sc. degree in Electrical Engineering from RWTH Aachen University, Germany. Currently, he is working as a research associate and Ph.D. student at Medical Information Technology, RWTH Aachen University. His research interests include bio-impedance spectroscopy and electrical impedance tomography.

3D simulation to analyse the effect of thrombosis on bioimpedance spectroscopy

Stephan DAHLMANNS[1], Leon GELLER-HARTUNG[1]

[1]Medical Information Technology, RWTH Aachen University, Pauwelsstr. 20, 52074 Aachen, Germany

dahlmanns@hia.rwth-aachen.de

Abstract. *Patients who suffer from circulatory disorders have a high risk of developing diseases in the lower leg, namely edema, thrombosis and chronic wounds. If bioimpedance spectroscopy (BIS) emerges suitable to detect these pathologies, it could be implemented in bandages or compression stockings to enable continuous monitoring and therefore improve medical treatment significantly. In this paper we analyse the effect of deep vein thrombosis on BIS using a 3D finite element (FEM) model of the dielectric properties of the human lower leg. Blood clot is modelled adopting permittivity and conductivity values for blood given in literature. Simulations are performed for a healthy leg and a leg where the vena tibialis is filled with blood clot. Results indicate that significant increases in resistance and reactance are detected by BIS, suggesting that deep vein thrombosis monitoring is possible using BIS technology.*

Keywords

3D FEM, Simulation, Bioimpedance Spectroscopy

1. Introduction

BIS is a commonly utilised monitoring technology to measure whole body properties in medicine [1]. Examples are body fat scales and the detection of muscle mass of geriatric patients. In recently published research, BIS is applied to investigate local changes in tissue properties like the healing of wounds [2, 3]. For further analysis and quantification, a 3D FEM model of the human lower leg was developed in [4]. Utilizing this model, tissue properties are modified and BIS is simulated. The simulation scenario introduced in this paper is the modulation of deep vein thrombosis by clogging the vena tibialis. Clogging of blood was accomplished by adjusting the dielectric properties inside the vein to simulate blood clots.

2. Methods

The 3D FEM model of the lower leg is based on anatomical data and composed of skin, subcutaneous fat, muscle, blood and bone. The dielectric properties of these tissues are derived from the Gabriel database [5]. As no values are available for blood clot, blood data are adopted and adjusted. Applying the values from Gabriel, the complex permittivity $\underline{\epsilon}$ for all tissues is calculated with the following Eq. 1 derived from Raicu [6],

$$\underline{\epsilon}(\omega) = \epsilon_\infty + \frac{\sigma'}{j\omega\epsilon_0} + \sum_{i=1}^{4} \frac{\Delta\epsilon_i}{1 + (j\omega\tau_i)^{1-\alpha_i}} \,. \tag{1}$$

where ω is the angular frequency, ϵ_0 vacuum permittivity, ϵ_∞, σ', ϵ_i, τ_i and α_i tissue properties. For a healthy lower leg, all necessary parameter values can be found in literature. For the creation of 3D FEMs, the permittivity $\epsilon(\omega)$ and conductivity $\sigma(\omega)$ needs to be calculated using

$$\epsilon(\omega) = \Re(\underline{\epsilon}(\omega)) \tag{2}$$

and

$$\sigma(\omega) = -\Im(\underline{\epsilon}(\omega)) \cdot \epsilon_0 \cdot \omega \,. \tag{3}$$

For the modelling of blood clot, its physiological properties are compared with blood. In a blood clot, cells are accumulated densely with few extracellular water. Assuming the Cole-Cole model of biological tissues [7], this reduction in extracellular water results in decreased conductivity, $\sigma_{\text{blood clot}} < \sigma_{\text{blood}}$. Subsequently, considering Eq. 1 and 3, blood clot can be modelled by adopting all blood parameters and then reducing σ' significantly. In this work it was set to

$$\sigma'_{\text{blood clot}} = \frac{1}{10} \cdot \sigma'_{\text{blood}} \,. \tag{4}$$

For the simulation of a deep vein thrombosis, the vena tibialis is filled with the modulated blood clot. This depicts a severe thrombosis, as in minor cases only parts of the vein are affected. All remaining blood vessels and tissues like muscle, subcutaneous fat and skin are not modified. The electrodes for current injection are modulated as well and placed on top of the skin layer approximately 14 cm apart, the voltage drop is calculated at two points between the input electrodes 12 cm apart. Simulations were performed for 30 frequencies logarithmically distributed between 1 kHz and 2 MHz for the lower leg with deep vein thrombosis as well as for the healthy leg.

3. Results and Discussion

The conductivity and permittivity of blood clot, blood and muscle are displayed in Fig. 1 in the range of the β-dispersion of biological tissues.

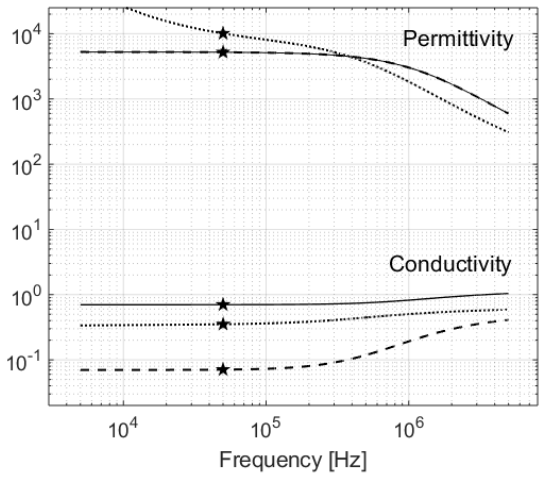


Fig. 1. Relative permittivity ϵ [n.u.] and conductivity σ [S/m] of blood (solid), blood clot (dashed) and muscle (dotted).

Blood and blood clot only differ in σ', therefore permittivity is identical for both tissues, assuming that cell properties are not altered when blood clots form. In contrast, conductivity is significantly lower for blood clot. At a common excitation frequency for the monitoring of tissue properties of 50 kHz, conductivity is reduced from 0.7008 S/m to 0.0708 S/m. Blood clot is less conductive than muscle, indicating that in BIS the current paths will shift from the normally high conductive blood vessels to muscle tissue. There, permittivity is larger for frequencies smaller 300 kHz ($+94.1\,\%$ at 50 kHz), hence an increase in reactance is to be expected for the simulated impedances.

The simulated results for the resistances and reactances monitored by BIS are depicted in Fig. 2.

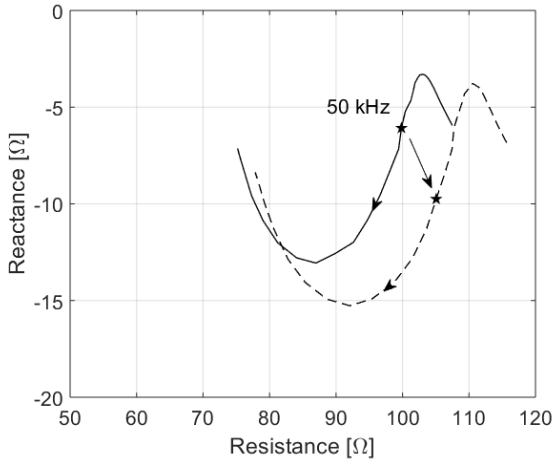


Fig. 2. Simulated BIS for the healthy lower leg (solid) and for the leg with thrombosis in the vena tibialis (dashed).

The computed values are connected with a solid line for the healthy leg and a dashed line for thrombosis. The resulting curves have a shape similar to a complex Cole-Cole plot and are in good agreement with literature values. The curve of the pathological leg shows larger resistances than the healthy leg. At 50 kHz, resistance increases by $5.3\,\%$ from $99.84\,\Omega$ to $105.10\,\Omega$. Additionally, a significant increase in reactance is visible. It increments by $60.4\,\%$ from $-6.08\,\Omega$ to $-9.75\,\Omega$. Analysing current densities in the 3D FEM model shows that current shifts from blood vessels to muscle tissue in the leg with thrombosis. As in our modulation the conductivity of blood clot is smaller than of muscle, the ratio of current flowing through the vena tibialis decreases. This leads to an increased proportion of muscle tissue information in the global BIS signal and a decrease in blood information. The increase in reactance for smaller frequencies illustrates this effect in particular. Absolute impedance increases by $5.58\,\Omega$ from $100.02\,\Omega$ to $105.6\,\Omega$, a variance detectable by BIS devices available on the market.

Our simulations show that BIS is a potential monitoring technique for thrombosis detection. In further studies, this effect has to be combined with subsequent changes in tissue properties like edema formation. This can be accomplished by changing subcutaneous fat properties in the model. Furthermore, the development of thrombosis over a time span of hours should be taken into consideration. This would allow for the investigation of BIS as an early warning system for thrombosis. Concluding, applying our models, guidelines for the analysis of leg status based on BIS monitoring could be developed in future work.

Acknowledgements

Research was supervised by Dr.-Ing. Daniel Teichmann and Prof. Dr.-Ing. Dr. med. Dr. h.c. (CTU Prague) Steffen Leonhardt, RWTH Aachen University.

References

[1] IVORRA, A. Bioimpedance monitoring for physicians: an overview. *Centre Nacional de Microelectrònica Biomedical Applications Group*, 2003, vol. 11, p.17.

[2] SWISHER, L., et al. Impedance sensing device enables early detection of pressure ulcers in vivo. *Nature Communications*, 2015.

[3] LUKASKI, H., et al. Bioelectrical impedance assessment of wound healing. *Journal of Diabetes Science and Technology*, vol. 6, 2012.

[4] DAHLMANNS, S., et al. 3D bioimpedance model of the human lower leg. *11th International Workshop on Impedance Spectroscopy*, 2018.

[5] GABRIEL, S., et al. The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues. *Physics in Medicine and Biology*, 1996, vol. 41, pp. 2271-2293.

[6] RAICU, V. Dielectric dispersion of biological matter: Model combining Debye-type and 'universal' responses. *Physical Review E60*, 1999, vol. 4, pp. 4677-4680.

[7] COLE, K. S., et al. Dispersion and absorption in dielectrics - alternating current characteristics. *The Journal of Chemical Physics*, 1941, vol. 9, pp 341–351.

Testbench to model cardiac volume changes

Leonie KORN[1], Max RUMPF[1]

[1]Medical Information Technology, RWTH Aachen University, Pauwelsstr. 20, 52074 Aachen, Germany

korn@hia.rwth-aachen.de

Abstract. *Ventricular Assist Devices (VADs) are used to treat patients with cardiogenic shock. However, there is still no consistent way to set pump speed. In order to control the pump speed of a VAD under consideration of the actual needs of the patient, it is necessary to measure and analyse relevant cardiac parameters. The positioning of minimal-invasive VADs in the left ventricle and the aorta can be utilized to apply impedance volumetry to measure left ventricular volume. In this work we present the development of a test bench to simulate changes in ventricular volume. The stroke volume of a ventricle phantom is dynamically controlled in a tank and thus offers the possibility to carry out in-vitro tests of impedance volume measuring systems. The test bench contains a linear motor controlled by a PID controller that moves three syringes. A dSPACE real-time system that connects all peripherals via the CAN bus interface serves as the higher-level user interface.*

Keywords

Volume change, Cardiac volumetry, Testbench

1. Introduction

People suffering from heart disease by 2015 numbered more than 400 million [1]. A dysfunction of the ventricles leads to a cardiogenic shock (CS), meaning insufficient blood flow into the organs. For the improvement of blood circulation, VADs are used to treat CS.

So far, setting the pump speed in VAD therapy is still a challenge, as an adequate reference is missing. Adapting to the individual needs of the patient is to be achieved by so-called physiological control. With this strategy, the physiological state of the patient is controlled instead of setting a fixed pump speed [2]. Minimal-invasive VADs are advantageously placed in the left ventricle and the aorta to measure parameters that help indicate the patient's physiological status attaching suitable sensors.

Placing electrodes on a VAD enables the observation of the electrical behaviour of the surrounding blood cavity and heart muscle tissue. It was observed by Baan et al. [3] that there is a correlation of impedance measurements in the ventricle and ventricular volume.

For further analysis of this correlation, customized heart phantoms to model anatomical and electrical behaviour of heart muscle tissue were developed in [4]. In this paper, a test bench to actuate these heart phantoms is presented. This allows for the optimization of the measuring system in a reproducible and cost-effective environment.

2. Hardware and Control of Testbench

The schematic of the testbench is shown in Fig. 1. The centrepiece of the testbench are three syringes, which are actuated by the linear motor DSAK4-24-50-300-IP54 (Drive-System Europe Ltd., Werther, Germany) to displace water in the pipes and the tank. The tank containing of

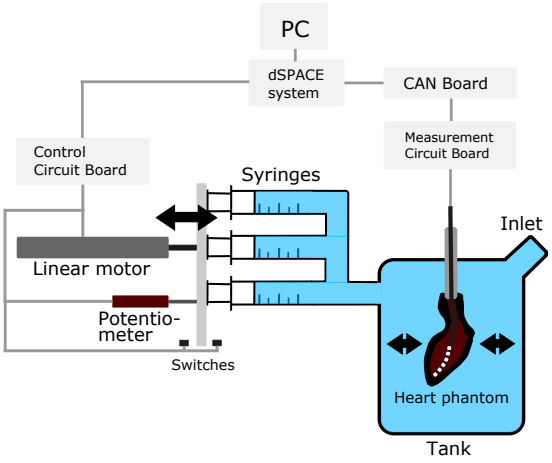


Fig. 1. Schematic of the testbench. Including the impedance measurement catheter and the heart phantom.

approx. 18 L water. The displacement of water in the tank leads to the displacement of fluid (e.g. saline solution) inside the heart phantom, as the tank is a closed system. The position of the linear motor is determined by a linear potentiometer. The displaced volume V in ml can then be calculated by measuring the distance d covered by the linear motor (see Eqn. 1). Depending on the geometry and the number of syringes selected, the factor to convert the distance into volume displacement has to be adjusted. In the presented case it is $K_{\text{ml:mm}} = 5.4\,\frac{\text{ml}}{\text{mm}}$.

$$V = d \cdot K_{\text{ml:mm}} \tag{1}$$

The Control Circuit Board is designed in a way that all functions required to drive the motor are concentrated on one board. Data and control parameters are exchanged with the

higher-level dSPACE system via the CAN interface. The microcontroller ATMega32M1 (Microchip Technology Inc., Chandler, USA) is used to control the H-bridge, the cooling fan and communicates via the CAN-Transceiver. The motor current of approx. 4.5 A is provided through the VNH5019A (STMicroelectronics, Amsterdam, Netherland) that includes a full H-bridge with an input voltage up to 24 V. To monitor and control the system, the microcontroller's ADC ports are used to measure motor current, system voltage, system temperature and motor position. Furthermore, it must be ensured that if the mechanical limits of the construction are exceeded, no damage is caused to the system. For this purpose, a safety circuit has been developed that can interrupt the power supply to the motor.
The electromechanical system was simulated with Simulink (The MathWorks Inc., Natick, USA) to better understand its behaviour and to tune and test the controller. Due to the presence of conjugated complex poles, the system is able to oscillate. In real application, the system is strongly damped by the large inertia of the piston tube and the gear of the motor as well as by friction. Consequently, a non-oscillating step response, so-called creep behaviour, was observed. For the test bench, a PID controller was chosen to drive the motor into a certain position. The requirements are high dynamics while allowing the lowest possible overshoot. The initial parameters of the PID controller are determined by analysing the step response of the system in simulation. These parameters are then used on the testbench as first control values and empirically adjusted. Additionally, an anti-windup is implemented to protect the motor against damaging input voltages. As an overall control system the MicroAutoBox II (dSPACE GmbH, Paderborn, Germany) is used to assign the set points for the motor position based on the desired heart volume changes. All parameters can be adapted in real-time and all measured quantities are displayed at the developed user interface.

3. Results and Discussion

For test measurements, the tank and all pipes were filled with water, sealed and vented. For instance, cardiac volume changes between 80 ml and 120 ml with a frequency of $f = 30\frac{\text{l}}{\text{min}}$ are reproduced by the testbench. The given volume set point and resulting actual volume are depicted in Fig. 2. There is no overshooting of the actual volume and the peak-to-peak values are successfully reached. The cross-correlation of both signals is 0.992 with a time delay of 100 ms. The time difference is due to a high integral term of the PID controller, but is necessary for the high accuracy of reaching a certain set point. Whereas, the time delay has no influence on the actual function of the test bench.

4. Conclusion

The in vitro test-bench presented is capable of dynamically adjusting the cardiac stroke volume of heart phantoms.

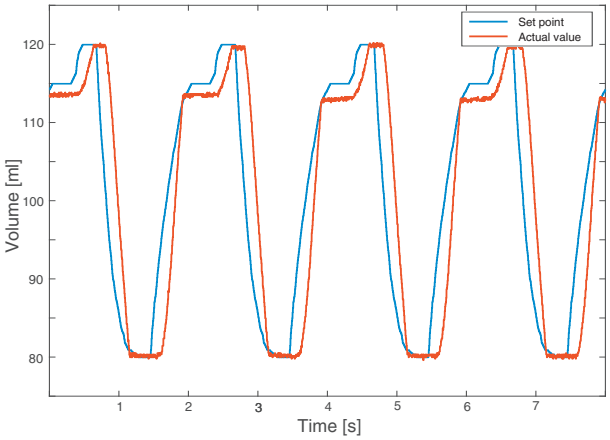


Fig. 2. Comparison between set point and actual value of volume changes in the testbench.

It was analysed in simulation to determine its most suitable control strategy. The selected hardware allows the monitoring and controlling of the entire system while guaranteeing a high level of reliability. The overall system convinces with its user-friendly interface and structured technical implementation. Therefore, it can be used to test and optimize measuring techniques that estimate ventricular volume. In the future, a variety of heart phantoms can be plugged into the tank and anatomical and electrical properties can be easily changed.
The limited dynamic range of the linear motor prevents very accurate cardiac volume curves for heart rates beyond $f = 60\frac{\text{l}}{\text{min}}$. In this case, sinusoidal or triangular volume changes can be applied so that a steep increase in volume can be avoided. The test bench is currently redesigned replacing plastic syringes with glass syringes to further increase the robustness of the system.

Acknowledgements

The presented research benefited from the input of Dr.-Ing. Marian Walter and Univ.-Prof. Dr.-Ing. Dr. med. Dr. h.c. Steffen Leonhardt, RWTH Aachen University.

References

[1] ROTH, G., JOHNSON, C., ABAJOBIR, A., *et al.*, Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015, *Journal of the American College of Cardiology*, 2017, p. 23715.

[2] ALOMARI, A., SAVKIN, A., STEVENS, M., , *et al.*, Developments in control systems for rotary left ventricular assist devices for heart failure patients: a review, *Physiological measurement*, 2012, vol. 34, no. 1, p. R1.

[3] BAAN, J., VAN DER VELDE, E., DE BRUIN, H., *et al.*, Continuous measurement of left ventricular volume in animals and humans by conductance catheter., *Circulation*, 1984, vol. 70, no. 5, pp. 812–823.

[4] KORN, L., LYRA, S., RÜSCHEN, D., PUGOVKIN, A., TELYSHEV, D., LEONHARDT, S., WALTER, M., Heart phantom with electrical properties of heart muscle tissue, *Current Directions in Biomedical Engineering*, 2018, vol. 4, no. 1, pp. 97–100.

Organic LED panels for pulse rate measurement using photoplethysmography imaging

Simon LYRA[1], Michael PAUL[1]

[1]Medical Information Technology, Helmholtz Institute for Biomedical Engineering,
RWTH Aachen University, 52074 Aachen, Germany

lyra@hia.rwth-aachen.de

Abstract. *Photoplethysmography Imaging describes a camera-based measurement technique that provides a method to acquire several vital signs such as heart rate and breathing rate without the need of direct skin contact. The intensity variations can be measured by using standard video cameras and suitable light sources. As of now, there is no established standard for light sources during the measurement, thus the results of different research groups are hardly comparable, because the modality is dependent on the interaction of light with the skin. In this paper, we would like to introduce a recording system with a novel illumination concept for the measurement technique, which uses organic light-emitting diodes (OLEDs). The recording system will be validated in a proof-of-concept study and the advantages of using OLED panels compared to conventional LEDs are pointed out.*

Keywords

Photoplethysmography Imaging, PPGI, remote vital signs, organic light-emitting diodes, OLED, medical imaging

1. Introduction

In recent years, the possibility of contactless vital parameter acquisition has been investigated by many research groups worldwide. The technique Photoplethysmography Imaging (PPGI) is the most common unobtrusive method to measure cardiovascular and respiratory processes where no direct skin contact is needed [1]. While respiratory signals can be measured by motion and modulation in the PPGI signal, the cardiovascular parameters can additionally be determined from variations in the intensity of backscattered and reflected light on the skin similar to contact-based PPG. These pulsatile variations are caused by the changing blood volumes in vessels during the heart cycle, which effects the pixel intensity on the sensor [1, 2].

In general, it is not entirely understood how the PPGI signal is generated. Besides the theory of direct measurement of changing blood vessel cross-sections, Kamshilin et al. and Sidorov et al. proposed an alternative theory based on the assumption that visible light will not penetrate down to pulsating arteries. They assumed a cyclic deformation of tissue in the dermis due to the pressure changes in larger arteries, which influences the capillary density and therefore the measured signal [3, 4]. Regardless of the underlying theory, the illumination of the measurement volume is crucial for a proper signal. When looking at the light sources used by research groups for PPGI measurements, a large variability can be observed: the sources range from ambient light, i.e. a mixture of sunlight and fluorescent tubes, to very specialized illumination schemes based on high-power quartz-halogen sources [5]. Depending on the target area of the measurement and the subsequent algorithm, ceiling lighting, light bulbs, RGB sources, infrared lighting or simple daylight were used [6, 7, 8].

One of the most advanced options is the use of high-output LEDs. These efficient and compact light sources are available in a variety of wavelengths and thus can be matched to the measurement scenario. However, this form of illumination has some disadvantages: due to the spotlight effect, the LEDs need to be combined with light guides and diffusors. In addition, this type of lamp always generates waste heat that can influence the measurement results, so when a LED panel is used, a heatsink must be considered.

Therefore, we would like to present a measurement system with OLED panels. We show that these are capable of providing sufficient amount of diffuse light for PPGI measurement and to compensate the disadvantages of LED technology. Further, to achieve high degree of homogeneity, the lighting panels are specially arranged. By using several light sources from different directions, it is possible to reduce shadows on the measurement object, thus the signal quality could be improved. Additionally, using the technology of OLED lighting, the overall system becomes more compact as these panels are surface emitter and not point sources, thus additional light guides and diffusors are not necessary.

2. Materials and Methods

Here, the used hardware and the measurement setup are introduced. A brief presentation of the sensor modality and the technology of OLEDs used in the setup is given.

2.1. Hardware

Camera Setup

The measurement setup is equipped with an industrial high-performance complementary metal-oxide-semiconductor (CMOS) monochrome camera of type Grasshopper 3 GS3-U3-23S6M-C (FLIR Integrated Imaging Solutions, Inc.). In 2017, Paul et. al. used several of these cameras for a multispectral setup to determine vital signs with PPGI [9]. Using the integrated IMX174 sensor (Sony Corporation), it is possible to perform measurements with a maximum resolution of 1920×1200 pixels up to $163\,\mathrm{Hz}$ with a global shutter. Furthermore, the provided USB3 interface allows high transfer rates. For PPGI measurements, a fixed focal length lens of type Fujinon CF12.5HA-1 (Fujifilm Holdings K.K.) was mounted on the camera. Additionally, for the proof-of-concept study in this paper a machine vision green interference filter of type Bi550-49 (Midwest Optical Systems) was used to concentrate on signals in one narrow wavelength range.

For the signal acquisition we used a six-core Intel Xeon E5-2620 V3 processor and 64GB DDR4 system memory. The Grasshopper 3 USB camera was connected to a controller card of type Point Grey ACC-01-1201 (FLIR) with a PCIe Express 2.0 x1 interface to enhance bandwidth. For recordings with high temporal and spatial resolution, several standard Solid-State-Drives (SSDs) and NVM Express SSDs were available.

Concept

Since Such, Acker and Blazek introduced PPGI in 1997, many research groups have evolved the method by using different cameras, light sources and advanced algorithms for signal extraction [10]. Especially the choice of intensity and wavelength of lighting represents an important factor for the signal quality. If the illumination intensity is too low, weak noise could already have a big influence on the signals. Furthermore, very bright light sources can lead to a saturation of the camera sensor. While single band lights are more robust compared to wide bands, white illumination is more natural for test subjects and has less impact on color perception in clinical scenarios such as operations.

OLED, like conventional LED, is a solid state lighting technology. Different from LED, the organic light panel is composed of $400\,\mathrm{nm}$ thin carbon based layers, which are sandwiched between two electrodes (see Fig. 1). Due to this structural form, the light is emitted from the whole surface, which provides a uniform illumination. The soft light casts

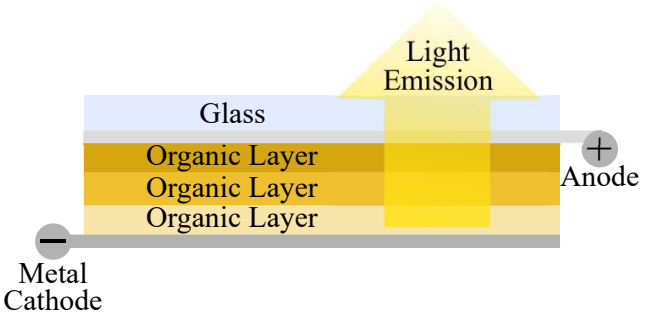


Fig. 1. Functional principle of an OLED panel.

less shadows and creates no glare. Another advantage is the fact that during light generation, neither UV nor NIR radiation is emitted [11]. These properties make the technology an appropriate light source for camera-based monitoring applications in the health sector. For the illumination of the camera setup, four OLED panels of type Keuka warmwhite (OLEDWorks) were used. The modules consist of a $45 \times 221\,\mathrm{mm}$ OLED panel and a driver encapsulated in a metal casing. The driver can dim the light between 5 and 100 percent by changing the electrical current in the OLED.

2.2. Trial Setup

The camera setup is given in Fig. 2. For this, a $3\,\mathrm{mm}$ thin aluminium base plate was prepared with mounting holes for different positions of the camera. Afterwards, the OLED panels were attached to it using self-designed 3D printed heat-resistant Acrylnitril-Butadien-Styrol (ABS) frames. While the inner two illuminate vertically from the plate, the angle of the outer two modules can be adjusted manually in order to determine the direction of light.

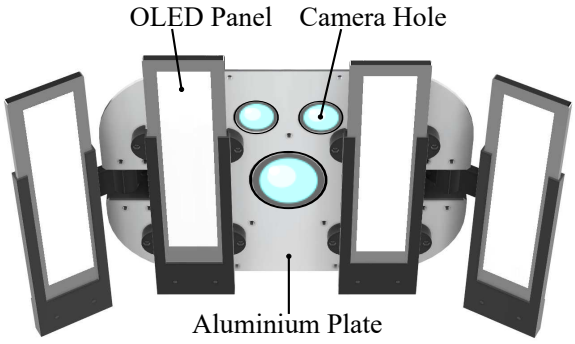


Fig. 2. Overview of the measurement system.

The illumination from several directions was implemented to use the advantages of diffuse front lighting combined with directional lighting from the side. This method can be used to reduce shading and compensate effects of ambient influences. The panels placed next to the camera can

light up the area with a shadow on one side of the object. This shadow can be reduced with the softer and less bright outer two OLEDs.

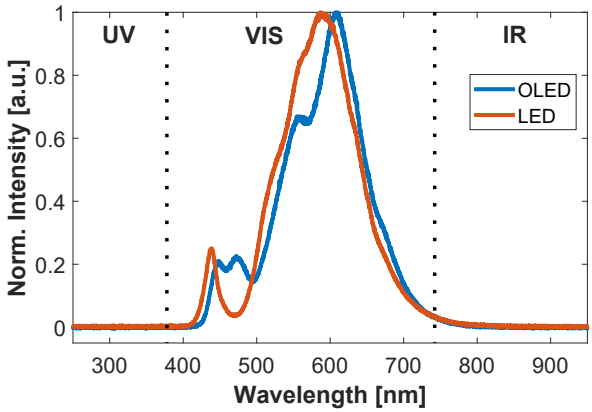


Fig. 3. Normalized spectra for OLED and LED.

For the proof-of-concept study, a subjects was asked to measure the PPGI signal on the back of the hand. For a first comparison, additional measurements with two LED spots of type S75-WHI-W (Smart Vision Lights) were performed. Each spot contains a 2×3 LED array. For diffuse light, optical filters of type Lee 416 and additionally a filter of type Lee 204 Full C.T. orange (LEE Filters) were applied. The second filter was used to convert the light from blueish to tungsten (orange) light.

To realize a first classification of the OLED light the emission of both modalities was measured with a compact CCD spectrometer of type Thorlabs CCS200/M (Thorlabs Inc.). The normalized spectra for $20\,\mathrm{ms}$ integration time and $10\,\mathrm{cm}$ distance to the sensor are shown in Fig. 3. It can be seen that the spectra are very similar, except higher intensities at wavelenghts of $450 - 500\,\mathrm{nm}$ and a drop at $580\,\mathrm{nm}$ are observable for OLEDs.

For the measurement the camera setup was positioned approx. $15\,\mathrm{cm}$ over the hand of the subject. Subsequently, the illumination was activated and repositioned to compensate shadows on the measurement volume. The back of the hand was recorded for $60\,\mathrm{s}$ with a sampling rate of $50\,\mathrm{Hz}$ using a resolution of 1920×1200 pixels. During the record, the subject was asked to hold the hand still. The reference was derived with $125\,\mathrm{Hz}$ sampling rate from a patient monitor of type MP70 (Philips) with a thumb-fixed PPG fingerclip.

3. Results

In Fig. 4 the calculated pulse rates from the measurement results of the proof-of-concept study are shown. To enable synchronicity related to the reference, the subject was asked to generate artifacts at the beginning and the end of the measurement, so the $10\,\mathrm{s}$ of samples at the on- and outset were neglected for signal analysis. To extract the pulse rate, a region of interest (570×570 pixels) was placed on the back of the hand and the pixel intensities in this region were averaged for every frame. The raw PPGI signals from the camera sensor were then filtered with a 2nd-order Chebyshev Type II bandpass filter with a lower stopband frequency of $0.3\,\mathrm{Hz}$ and a higher stopband frequency of $5\,\mathrm{Hz}$ using a stopband attenuation of $20\,\mathrm{dB}$. After filtering, the pulse rate could be derived by a peak detection algorithm. For the reference, just the peak detection was applied to the raw signal.

It can be observed that both OLEDs and LEDs enable a reliable pulse rate extraction using PPGI. In a first analysis, the OLED measurement system seems to be less sensitive to fast heart rate changes than the one with LEDs. This process can be observed during times of dynamic pulse rate

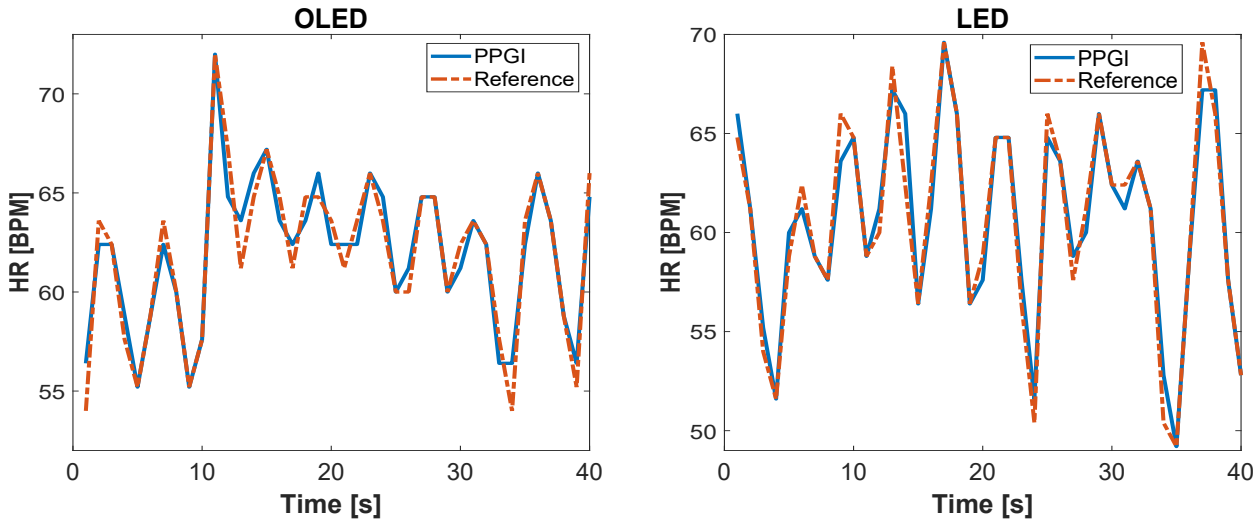


Fig.4. Results of the proof-of-concept PPGI study using OLED (left) and LED (right) illumination.

changes. Finally, the values for the RMSE can be calculated to $RMSE_{\mathrm{OLED}} = 1.106$ and $RMSE_{\mathrm{LED}} = 1.138$.

The low value for the RMSE for OLEDs show the suitability for PPGI measurements. Finally, the general possibility of contactless pulse measurement and the described advantages of this light source illustrate the great potential of OLED technology for applications in the medical technology sector.

4. Discussion and Conclusion

In this paper, we introduced a measurement setup for PPGI with OLED panels as novel illumination source. We performed measurements with the system and compared the signals with a PPG reference. The measurement results show that the used OLED panels are generally suitable as illumination source for the application in camera setups to derivate a PPGI signal. In the future, a larger PPGI study with a statistical analysis needs to be performed. It is hoped that our findings lead to an increased use of organic LEDs in PPGI measurement setups and in biomedical engineering in general due to the many advantages, which are provided by this light technology. Additionally, the use of direct and indirect illumination is recommended for all future PPGI measurements to compensate shadows on the skin and optimize signal qualities.

Acknowledgements

Research described in the paper was supervised by Prof. Dr.-Ing. Dr. med. Steffen Leonhardt and Dr.-Ing. Christoph Hoog Antink. Simon Lyra gratefully acknowledges financial support provided by the German Research Foundation (DFG), grant no. LE 817/32-1.

References

[1] HUELSBUSCH, M. Ein bildgestütztes, funktionelles Verfahren zur optoelektronischen Erfassung der Hautperfusion (A functional imaging technique for optoelectronic assessment of skin perfusion). Ph.D. thesis, RWTH Aachen, 2008.

[2] ZAUNSEDER, S., TRUMPP, A., WEDEKIND, D., MALBERG, H.. (2018). Cardiovascular assessment by imaging photoplethysmography - A review. Biomedical Engineering / Biomedizinische Technik. 63. 10.1515/bmt-2017-0119.

[3] KAMSHILIN AA, NIPPOLAINEN E, SIDOROV IS, VASILEV PV, EROFEEV NP, PODOLIAN NP, et al.; A new look at the essence of the imaging photoplethysmography. Sci Rep 2015;5:10494.

[4] SIDOROV IS, ROMASHKO RV, KOVAL VT, GINIATULLIN R, KAMSHILIN AA.; Origin of infrared light modulation in reflectance-mode photoplethysmography. PLoS One 2016;11:1–11.

[5] BLACKFORD, E.B., PIASECKI, A.M., ESTEPP J.R.; Measuring pulse rate variability using long-range, non-contact imaging photoplethysmography; Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, vol. 2016–Octob, pp. 3930–3936, 2016.

[6] VOGELS, T., VAN GASTEL, M., WANG, W., DE HAAN, G.; Fully-Automatic Camera-Based Pulse-Oximetry During Sleep; The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2018, pp. 1349-1357

[7] MOCO, A. V., STUIJK, S., DE HAAN, G. (2016). Skin inhomogeneity as a source of error in remote PPG-imaging. Biomedical optics express, 7(11), 4718-4733. doi:10.1364/BOE.7.004718

[8] WU, T., BLAZEK, V., Schmitt, H.J.; "Photoplethysmography imaging: a new noninvasive and noncontact method for mapping of the dermal perfusion changes", Proc. SPIE 4163, Optical Techniques and Instrumentation for the Measurement of Blood Composition, Structure, and Dynamics, (22 November 2000); doi: 10.1117/12.407646; https://doi.org/10.1117/12.407646

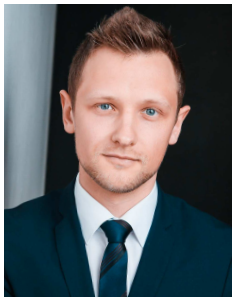
[9] PAUL M., VENEMA B., BLAZEK V., MÜHLSTEFF J., LEONHARDT S. (2018) A camera-based multispectral setup for remote vital signs assessment. In: Eskola H., Väisänen O., Viik J., Hyttinen J. (eds) EMBEC and NBC 2017. EMBEC 2017, NBC 2017. IFMBE Proceedings, vol 65. Springer, Singapore

[10] SUCH, O., ACKER, S., BLAZEK, V.. (1997). Mapped hemodynamic data acquisition by near infrared CCD imaging. 2. 637 - 639 vol.2. 10.1109/IEMBS.1997.757694.

[11] OLED Works (Rochester, NY, USA); URL: https://www.oledworks.com/oled-technology/; (04.03.19)

[12] BIRN, J., Digital Lighting and Rendering. In: New Riders digital series, 2000. ISBN 1562059548, 9781562059545.

About Authors...



Simon LYRA was born 1994 in Essen, Germany. In April 2018 he received the M.Sc. degree in Electrical Engineering with specialisation on Biomedical Engineering from the RWTH Aachen University, Germany. Currently he is working as a research associate and Ph.D. student at Medical Information Technology, Helmholtz-Institute for Biomedical Engineering at RWTH Aachen University. His research interests include remote vital parameter measurement with a focus in Photoplethysmography Imaging for the application in neonatal intensive care units.

Michael PAUL was born in Dormagen, Germany, in 1988. In 2014 he received his M.Sc. degree in Electrical Engineering with the major field of study in Biomedical Engineering from RWTH Aachen University, Germany. Since then, he is a Ph.D. student at Medical Information Technology, Helmholtz-Institute for Biomedical Engineering at RWTH Aachen University. His research interests are camera-based measurement techniques in healthcare applications with focus on continuous vital signs monitoring using Photoplethysmography Imaging and Infrared Thermography.

High frequency oscillation analysis in epileptic patients

Martin DOSTÁL[1]

[1] Dept. of Circuit Theory, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

dostam15@fel.cvut.cz

Abstract. *It is possible to cure some patients suffering from refractory epilepsy by performing resective surgery, in which a region of brain which causes the epilepsy is removed from the patient. The accuracy of the estimation of size and position of the region-to-be-resected could be improved by being able to infer the exact or approximate diagnosis of the type of the epilepsy that is being dealt with. This works tries to use the relationship between two biomarkers, HFO and IED to differentiate between FCD epilepsy types to provide valuable presurgical information which could improve the resective surgery outcome.*

Keywords

EEG, HDEEG, epilepsy, HFO, signal processing

1. Introduction

Some patients who suffer from refractory epilepsy (conventional medication does not work) can undergo resective surgery, during which a part of the brain tissue which is anticipated to be the cause of the epilepsy is removed [1].

The traditional biomarker, used for delineation of the region of the brain to be removed during the resective surgery, is Interictal Epileptiform Discharge (IED). IEDs are known to be fairly reliable in estimation of the area-to-be-resected, but oftentimes, they lack in specificity.

High Frequency Oscillations (HFOs) is a rather new putative biomarker of epileptic brain regions. HFOs do not have unified definition at the time of writing of this work, but authors usually agree, that it is an oscillatory activity in frequency band 80-500 Hz, with at least three consecutive oscillations which do have noticeably higher amplitude than the signal's background activity[2].

The goal of this work is to assess the HFO activity in the temporal vicinity of IEDs in the patient cohort of epileptic child patients and to try and find, if in some patients exists a temporal relation between IEDs and HFOs. The main hypothesis of this work is, that it is possible to sort patients into three groups (no HFO, HFO not temporarily locked to IED, HFO temporarily locked to IED) and that these three groups could correspond to some clinically interesting facts.

2. Data

The data used in this work are 128 channel HDEEG data recorded at sample rate of 2048 Hz in FN Motol hospital as a part of presurgical child patient evaluation. The cohort of patients used in this work consists of 25 child patients with presumably varied diagnoses.

3. Methods

To assess patient in term of HFO markers, it is first needed to answer the question whether the patient shows sufficient evidence of HFO existence. If the answer to the previous question is positive, then the temporal relationship between HFOs and IEDs needs to be addressed. These two stated problems are answered by the methods described in the following text.

3.1 HFO existence check

Significant presence of HFOs in the patients is assessed based on an averaged time-frequency scalogram (Morlet wavelet) across all IED events of one type (IEDs are clustered by their simultaneous activity on specific channels) in each patient. Such an averaged time-frequency event can be seen in Fig. 1

First, two energetic profiles are made, one is that of a spike and one of the signal's background. The spike profile is made by averaging time bins in the neighborhood of the spike position and the background profile is made by averaging time bins which are not located in spike time interval neither in border transients created by the wavelet transform. Difference of the spike and background energy profile (hereinafter referred as to "profile"), shows the "unusualness" of the signal in the spike time interval.

The value of the profile at the start of the HFO search interval (where the HFO is expected to occur) is set as a threshold value. The HFO search interval is then searched for peaks on the profile and the strongest peak is selected. If the selected peak is higher than the threshold set, it is labeled as a putative HFO of unknown prominence. The prominence is computed as an area of the peak which exceeds the threshold (can be seen in Fig. 2).

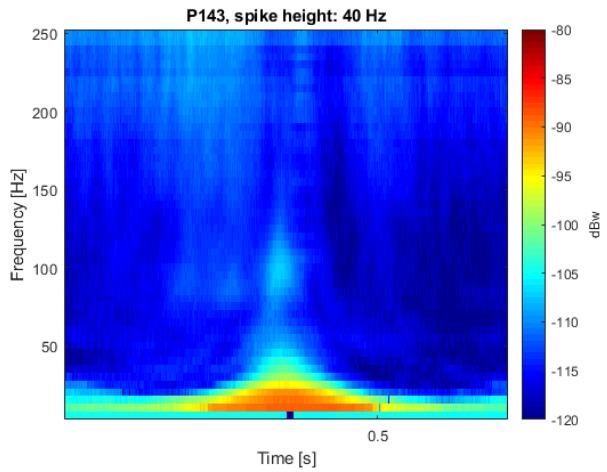


Fig. 1 Averaged scalogram of the IED events of one type in patient P143

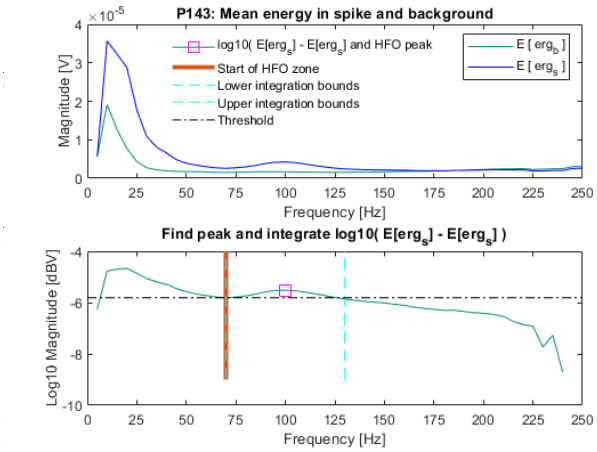


Fig. 2 Process of determining whether HFO is present or not

It is possible to use the HFO prominence to sort patients into two groups: HFO active and HFO inactive.

3.2 HFO – IED temporal relationship

To answer the question of temporal relationship between a time-frequency (scalogram) HFO detector was developed. The detector's core functionality is based upon the fact, that the phase information from the complex scalogram changes more slowly when HFO is present thanks to its similarity to the Morlet wavelet used in the wavelet transformation. The detections and their counts are depicted in Fig. 3.

4. Preliminary results and discussion

Thanks to the lack of clinically determined diagnosis, it is necessary to use some other means of reference, when assessing the results of this work. For the preliminary reference, human labeling done by my supervisor (experienced EEG analyst) and myself was aggregated to maximize specificity and suppress sensitivity. No statistical significance was found between the results given by the methodical approach and between the aggregated human labeled data (reference).

The reason for this statistically insignificant result might be a mistake in the methodical approach. The human detector does not detect weak HFOs, only strong ones. The total population of these strong HFOs is presumably quite small, so the averaging process of the events at the beginning of the methodical approach suppresses these human-detected events and thus creates discrepancy between the human and the machine detector.

As the majority of subjects from the patient cohort are awaiting surgery, which will enable the clinical diagnosis to be carried out, the reference diagnosis is not yet available, but will be in the future.

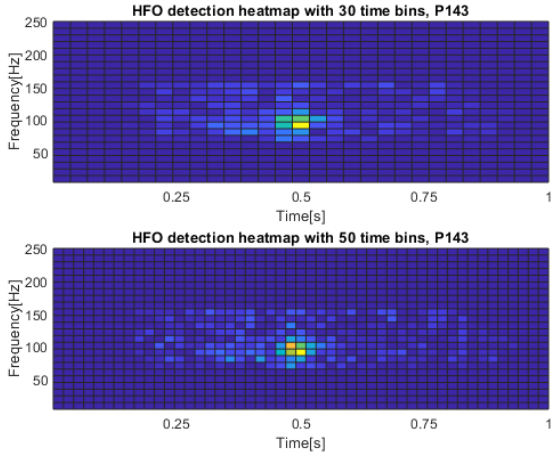


Fig. 3 Heatmaps of HFO distribution in time-frequency with various temporal resolution

Acknowledgements

Research described in the paper was supervised by Ing. Petr Ježdík PhD, FEE CTU in Prague and supported by the Czech Ministry of Healthcare under grant NV17-28427A.

References

[1] R. S. FISHER et al., "ILAE Official Report: A practical clinical definition of epilepsy," Epilepsia, vol. 55, no. 4, pp. 475–482, 2014.

[2] B. FRAUSCHER et al., "High-frequency oscillations: The state of clinical research," Epilepsia, vol. 58, no. 8, pp. 1316–1329, 2017.

About Author

Martin DOSTÁL was born in Prague, Czech Republic and is currently finishing his master's degree in electrical engineering at FEL ČVUT.

Remote breathing signal extraction using infrared thermography in a climatic chamber

Christoph WEISS[1], *Alexander KIRMAS*[2]

[1]Medical Information Technology, Helmholtz-Institute for Biomedical Engineering at RWTH Aachen University
[2]Institute for Automotive Engineering at RWTH Aachen University

christoph.weiss@rwth-aachen.de, alexander.kirmas@ika.rwth-aachen.de

Abstract. *Being an indicator of physiological distress, the importance of breathing rate (BR) monitoring is increasing. As already shown in previous publications, unobtrusive BR monitoring is possible with infrared thermography (IRT). While the typical ambient temperatures in the previous measurements were around 22 °C, the aspect of varying ambient temperatures and their influence on BR extraction from IRT has to be examined for robust extraction algorithms. Therefore, this paper describes a study using IRT in a climatic chamber. While recording 44 subjects with a far infrared camera in a sitting position without movement in eight different clothing and temperature variations, a reference measurement system, a chest strap, was set up as well for recording the reference breathing signal. Using the extraction method described by Murthy & Pavlidis, we show the breathing signals under different environmental conditions.*

Keywords

infrared thermography, breathing extraction, climate chamber, FFT.

1. Introduction

Monitoring health by evaluating vital signs is a common task in clinical environment. One parameter often neglected is the breathing rate. Usually, this parameter is acquired by contact based procedures like spirometry through airflow measurement or by length changes of a chest strap due to thoracic movement [3].
Using unobtrusive techniques to extract vital signs is a core area in medical engineering. For breathing rate detection, different methods have been shown based on radar or cameras. While camera based approaches can differ in the used wavelength, we concentrate on a far infrared wavelength camera. In contrast to radar or near infrared measurements, the thermal infrared is completely passive and does not rely on lighting conditions. Describing a virtual thermistor Murthy & Pavlidis [1] demonstrated a contact-free method by utilizing the exhaled airstream heated up by the body and visible in the thermal infrared camera stream. To extend the possibility of this acquisition modality also shown by [2] at 22 °C we acquired recordings in a climate chamber at different ambient temperatures.

2. Test set-up

The test set-up was built in a climate chamber. This chamber can be adjusted to different temperatures (in a range of -20 °C to 40 °C) with a lead time of about 15 minutes.
In total 44 subjects were sitting on a chair in front of a thermal infrared camera, InfraTec VarioCam HD head 820 / 30 mm running at 15 fps. The camera was placed in a distance of about 1.2 m, resulting in an instantaneous field of view of 0.68 mm/px. Varying the conditions all subjects were recorded eight times for about one minute as depicted in Tab. 1. The ventilation scenario was obtained with commercially available fans (see Fig. 1), one in height of the chest, the other blowing up from the feet. For every temperature change the subjects had 15 minutes of acclimatization time.

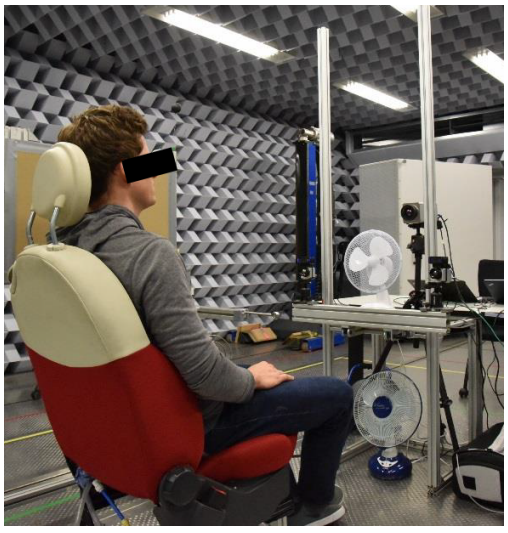


Fig. 1. Recording set-up in the climatic chamber. IRT camera and fans are placed in front of the subject.

Referencing of the recordings was done by data acquisition with Somnolab 2 PSG (Weinmann GmbH), a polysomnographic device. Data included electrocardiogram (3 electrodes, 256 Hz), photoplethysmograph (finger clip, 50 Hz) and breathing signal by a chest strap with a sampling rate of 32 Hz.

3. Results

Extraction of the breathing signal from the camera stream could be achieved by the method shown by Murthy & Pavlidis [1]. During inspiration and expiration, the air is heated up inside the body and is visible in the thermal camera stream when exhaling. Using a fixed region of interest (ROI) the average ROI size of extraction was 60 px by 40 px, resulting in 40 mm by 27 mm of facial area. The extracted signal and the reference signal were both filtered by FIR bandpass with cut-off frequencies 0.05 Hz and 0.4 Hz. As visible in Fig. 2, the peaks of reference and camera-extracted signals are alternating due to the different acquisition techniques. While breathing out the belt will tighten, leading to a local minimum. In contrast, the heated airstream of the nostril region yields a local maximum. All shown signals are normalized after filtering to a range of -1 to 1. Using Fast Fourier Transform (FFT) the BR could be detected by selecting the most significant component in the amplitude spectrum. As shown in Tab. 1, the BR extraction in all variations of temperature and clothing was possible and the difference to the reference is negligible.

Temperature	T	P	J	V	BR_{C}	BR_{Ref}
10 °C		**x**			0.26	0.26
10 °C			**x**		0.19	0.21
20 °C	**x**				0.26	0.26
20 °C		**x**		**x**	0.2	0.2
20 °C		**x**			0.23	0.2
30 °C	**x**				0.23	0.22
30 °C		**x**		**x**	0.21	0.22
30 °C		**x**			0.26	0.25

Tab. 1. Extracted frequencies from the camera stream (BR_{C}) and reference frequencies (BR_{Ref}) for subject 44 in all recording variations (t-shirt (T), pullover (P), jacket (J), ventilation (V)) in the climate chamber.

4. Conclusion

We demonstrated the possibility of obtaining the BR under a set of temperature (and clothing) variations. In our test set-up, we showed the remote extraction of the breathing rate from a thermal infrared camera in eight different scenarios and validated the extracted signals against a reference signal. Further studies will have to show if different temperatures have an influence on breathing strength or breathing patterns.

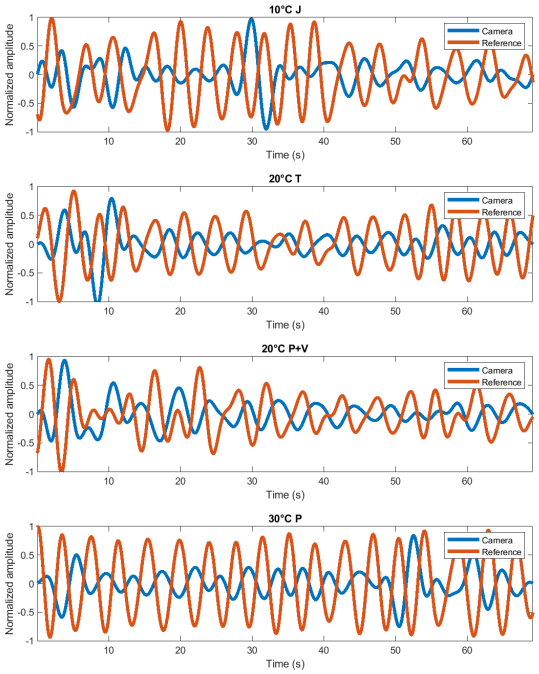


Fig. 2. Reference (red) and camera-extracted (blue) breathing signals of subject 44 at 10 °C (Jacket), 20 °C (T-Shirt), 20 °C (Pullover and ventilation), 30 °C (Pullover). The alternating peaks are due to the different acquisition techniques (see section 3).

Acknowledgements

Research described in this paper was supervised by Dr. Marian Walter and Prof. Dr.-Ing. Dr. med. Dr. h.c. Steffen Leonhardt, RWTH Aachen University, Germany. We also want to thank Damian Backes M.Sc. and the Institute for Automotive Engineering at RWTH Aachen University for giving the opportunity to use the climate chamber.

References

[1] MURTHY, R.; PAVLIDIS, I.: NON-CONTACT MONITORING OF BREATHING FUNCTION USING INFRARED IMAGING, 2005

[2] PEREIRA, C. B. et al.: Remote monitoring of breathing dynamics using infrared thermography, Biomedical optics express vol. 6,11 4378-94., 2015, doi:10.1364/BOE.6.004378

[3] AL-KHALIDI, F. Q.; SAATCHI, R.; BURKE, D.; ELPHIK, H.; TAN, S.: Respiration rate monitoring methods: a review, Pediatric pulmonology 46,6 523-9., 2011, doi:10.1002/ppul.21416

About Authors...

Christoph WEISS was born in Bonn, Germany, in 1988. He is currently enrolled for materials engineering.

Alexander KIRMAS was born in Düsseldorf , Germany, in 1987. He is currently enrolled as a PhD student.

Feedforward Compensation of Physical Activity in Insulin Therapy for Type 1 Diabetes using Oxygen Uptake

Bianca ROMANSKI[1], *Carl-Friedrich BENNER*[1]

[1] Medical Information Technology, RWTH Aachen University, Pauwelsstr. 20, 52074 Aachen, Germany

bianca@romanski-bs.de, benner@hia.rwth-aachen.de

Abstract. *The blood glucose control in patients with diabetes mellitus type 1 is achieved by an administration of insulin, of which the correct dosage is essential. A new therapy approach is the automated insulin delivery utilizing a continuous glucose measurement. Due to its impact on the glucoregulatory system, physical activity is one of the major hurdles of automated insulin delivery. In this paper, feedforward compensators for physical activity using an oxygen uptake measurement are developed. An adapted version of the Sorensen glucoregulatory model is used for both the design and the in silico evaluation of the compensators. The design follows the design rules by Guzmán et al. The results show the potential of feedforward compensators to improve blood glucose control during physical activity.*

Keywords

glucoregulatory model, physical activity, feedforward compensation, type 1 diabetes, closed-loop

1. Introduction

Diabetes mellitus type 1 is a metabolic disease in which the patient's pancreas lacks the production of the blood glucose (BG) lowering hormone insulin. The therapy for type 1 diabetes is an exogenous injection of insulin, aiming for near-normal BG levels (80-120 mg/dl). An adequate insulin dose is essential to prevent hypo- and hyperglycemia [1]. Multiple disturbances (e.g. exercise, food intake) and the delayed effect of subcutaneously injected insulin impede BG control. The development of systems for continuous glucose monitoring (CGM) enabled a paradigm shift in insulin therapy. The new approach is a so called Artificial Pancreas, which automatically adapts the insulin infusion rate (IIR) utilizing information received from the CGM [2].

Physical activity (PA) influences the glucoregulatory system significantly. The impact depends on type and duration of an activity and differs in individuals [3], which hinders an adequate adaption of the insulin therapy. Informing an Artificial Pancreas about PA enables an improved control compared to uninformed systems [4]. In this article, a feedforward compensator (FFC) utilizing an oxygen uptake measurement, i.e. detecting PA automatically and respecting its intensity, with the aim to improve BG control during PA is developed.

2. Methods

2.1 Glucoregulatory Model

For the design and evaluation process, a glucoregulatory model, which describes the effect of PA as a function of measurable input signals characterizing PA, is needed. An adapted version of the Sorensen [5] model with extensions by Lenart et al. [6] and Hernández-Ordoñez et al. [7] meets these requirements. The model describes glucose, insulin and glucagon concentrations in six different parts of the body. Distinguished compartments are brain, heart and lung, liver, gut, kidneys and periphery (i.e. muscle and fat tissue). In compartments with a non-negligible diffusion time, there is a further distinction between the vascular space and the interstitium. Multiple physiologic sources and sinks are modeled. Endogenous insulin secretion is removed to describe a patient with type 1 diabetes.

PA is characterized by the percentage of maximum oxygen uptake (PVO_2^{max}) and by the percentage of active muscle mass (PAMM). It should be noted that the input variable for the oxygen uptake can be discontinuous, as the dynamic evolution of PVO_2^{max} is included in the model. Modelled effects of PA on BG include an increase in peripheral glucose and insulin uptake, an increase in hepatic glucose production and an adjustment of the blood flow distribution. Both short- and long-term aerobic exercise can be described, as the depletion of the hepatic glycogen reservoir is modelled. The extended Sorensen model was further adapted. Amongst others, it was extended with models by Dalla Man et al. [8] to describe the delayed absorption of subcutaneously injected insulin and the delay between the BG and the CGM signal. The model used in this paper consists of 26 nonlinear differential equations.

2.2 Feedforward Compensator

The effect of PA is described in terms of PVO_2^{max} and PAMM. Therefore, these signals can be considered as disturbances and are possible inputs for a FFC. The design of the FFC follows the design rules by Guzmán et al. [9]. There are two issues when applying these rules to the glucoregulatory model.

Firstly, Guzmán et al. consider only one disturbance signal. One approach is to treat PVO_2^{max} and PAMM as independent disturbances by designing two FFCs and adding their outputs. This is problematic, as most of the modelled effects are influenced by both input signals, showing nonlinear interactions with each other. Moreover, the impact of PVO_2^{max} (i.e. the intensity) is larger than the impact of PAMM (i.e. the type of the aerobic exercise). Therefore, only PVO_2^{max} is considered as disturbance input signal. PAMM is assumed to be 25 % during the FFC design process.

Secondly, the prerequisite for using the design rules is that both the process (influence of changes in IIR on BG) and the disturbance (influence of changes in PVO_2^{max} on BG) transfer function are PT1 elements with delay (PT_1T_t), because the FFC transfer function is derived from the parameters of these PT_1T_t elements. As this is not the case, simplified substitutional transfer functions are derived from step responses of the model for the design process. A substitutional transfer function consisting of a single PT_1T_t element approximates the process transfer function appropriately (see Fig. 1). Multiple insulin step responses are investigated (steps from basal IIR to values between 0.2 and 2 times basal IIR) and mean values are used for the process transfer function (see Eq. 1).

$$G_{\text{Insulin}}(s) = \frac{-3.9}{1+116s} e^{-55s} \tag{1}$$

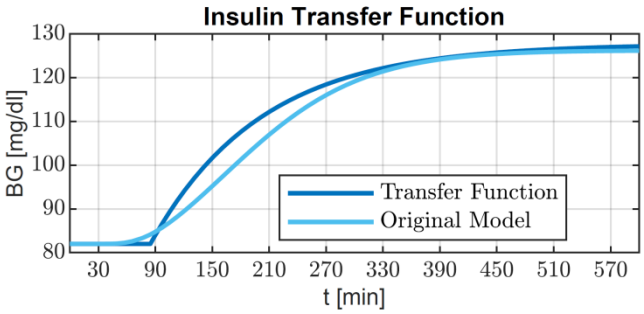


Fig. 1. Approximating the step response of the original model with a substitutional transfer function consisting of a single PT_1T_t element. IIR is halved at t = 30 min.

In contrast, the disturbance transfer function is not approximated satisfactorily by a single PT_1T_t element (see Fig. 2). To improve the approximation while still being able to use the design rules, two accumulated PT_1T_t elements with different delays are used (see Eq. 2). This means that the total PA disturbance is modeled as two individual disturbances depending on the same input signal (i.e. PVO_2^{max}). Here, the less delayed PT_1T_t element represents the general effect of PA, whilst the more delayed PT_1T_t element represents exhaustion due to the depletion of the hepatic glycogen reservoir. One FFC is derived for each PT_1T_t element using the design rules. The complete FFC is obtained by accumulating both FFCs. The substitutional disturbance transfer function is derived from a step response to PAMM = 25 % and PVO_2^{max} = 30 %.

$$G_{\text{Activity}}(s) = \frac{-0.34}{1+53s} e^{-7s} + \frac{-1.1}{1+88s} e^{-176s} \tag{2}$$

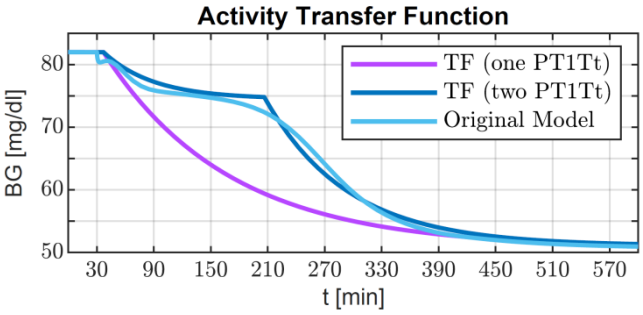


Fig. 2. Approximating the step response for a change in PVO_2^{max} and PAMM from the basal state, i.e. 8 % and 0 %, to 30 % and 25 % at t = 30 min. Comparison of the original model with substitutional transfer functions consisting of one or two PT_1T_t elements.

Guzmán et al. describe four different types of FFCs. Eq. 3 is the transfer function of a "static" FFC, Eq. 4 of a "static with delay" FFC, Eq. 5 of a "lead-lag" FFC and Eq. 6 of a "lead-lag with delay" FFC. In contrast to static FFCs, lead-lag FFCs consider the different time constants of insulin and activity transfer function. FFCs with delay take the delayed onset of the exhaustion effect into account, whereas FFCs without delay reduce IIR as if exhaustion already sets in at the beginning of PA.

$$G_{\text{S}}(s) = -0.089 - 0.275 = -0.364 \tag{3}$$

$$G_{\text{SD}}(s) = -0.089 - 0.275 e^{-122s} \tag{4}$$

$$G_{\text{L}}(s) = -0.089 \frac{1+116s}{1+53s} - 0.275 \frac{1+116s}{1+88s} \tag{5}$$

$$G_{\text{LD}}(s) = -0.089 \frac{1+116s}{1+53s} - 0.275 \frac{1+116s}{1+88s} e^{-122s} \tag{6}$$

2.3 Examined Scenario

FFCs are validated in both open and closed loop control using the glucoregulatory model. For closed loop control, the PID controller by Zimmermann [10] (K_P = -0.5, T_I = 240 min, T_D = 30 min) is used. The investigated PA is characterized by PVO_2^{max} = 60 % and PAMM = 25 % and lasts from t = 30 min to t = 300 min.

3. Results

Fig. 3(a) shows BG in open loop control. In all cases, there is a comparable decline in blood glucose within the first 30 min of PA. Continuation of the basal IIR results in a second period with a pronounced decline in BG at t = 180 min, yielding severe hypoglycemia. The use of FFCs with delay leads to near-constant BG values from t = 60 min to t = 210 min, i.e. after the initial decline of BG. Afterwards, BG rises.

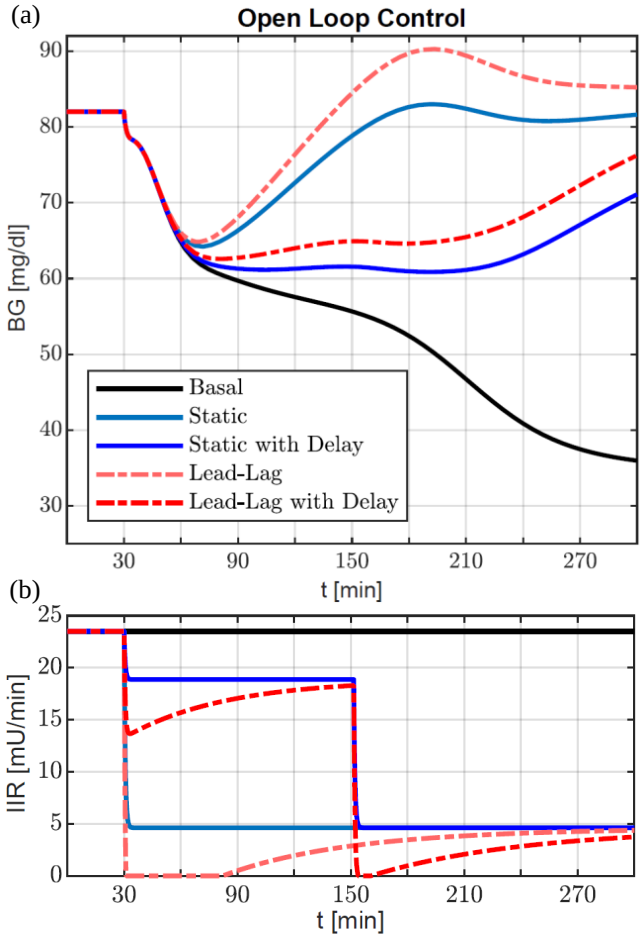


Fig. 3. (a) BG and (b) IIR during PA in open loop control using continuation of the basal IIR and one of the developed FFCs (static, static with delay, lead-lag, lead-lag with delay).

In contrast, FFCs without delay cause an increase after the initial decline till BG peaks at approximately t = 190 min at values higher than the basal BG. Tab. 1 summarizes resulting time in hypoglycemia (< 70 mg/dl), minimal BG and root-mean-square deviation (RMSD) from the basal state. All FFCs reduce time in hypoglycemia and RMSD and increase minimal BG. FFCs without delay cause larger changes of these values than FFCs with delay.

	Time in Hypoglycemia [min]		Minimal BG [mg/dl]		RMSD [mg/dl]	
	OL	CL	OL	CL	OL	CL
Basal	249	-	36	-	31.5	-
PID	-	249	-	62	-	15.5
S	57	47	64	65	8.1	7.8
S D	240	74	61	63	18.4	10.2
LL	47	46	65	65	8.3	8.1
LL D	203	61	63	64	15.4	9.1

Tab. 1. Time in hypoglycemia, minimal BG and RMSD from the basal BG in open loop (OL) and closed loop (CL) control using continuation of basal IIR, exclusive PID control and the developed FFCs (S: static, S D: static with delay, LL: lead-lag, LL D: lead-lag with delay).

Fig. 4(a) shows BG in closed loop control. Closing the loop with a PID controller counteracts the decline in BG. As can be seen in Tab. 1, RMSD is reduced and minimal BG is increased compared to continuation of basal IIR. However, the PID controller does not reduce time in hypoglycemia, as hypoglycemia persists. Combining a FFC and a PID controller leads to a reduction of time in hypoglycemia and RMSD and to an increase of minimal BG compared to exclusive PID control and to open loop control with the same FFC. Like in open loop control, FFCs without delay lead to a smaller time in hypoglycemia, a higher minimal BG and a lower RMSD than those with delay. The static FFC leads to the lowest RMSD, whereas the lead-lag FFC causes the lowest time in hypoglycemia.

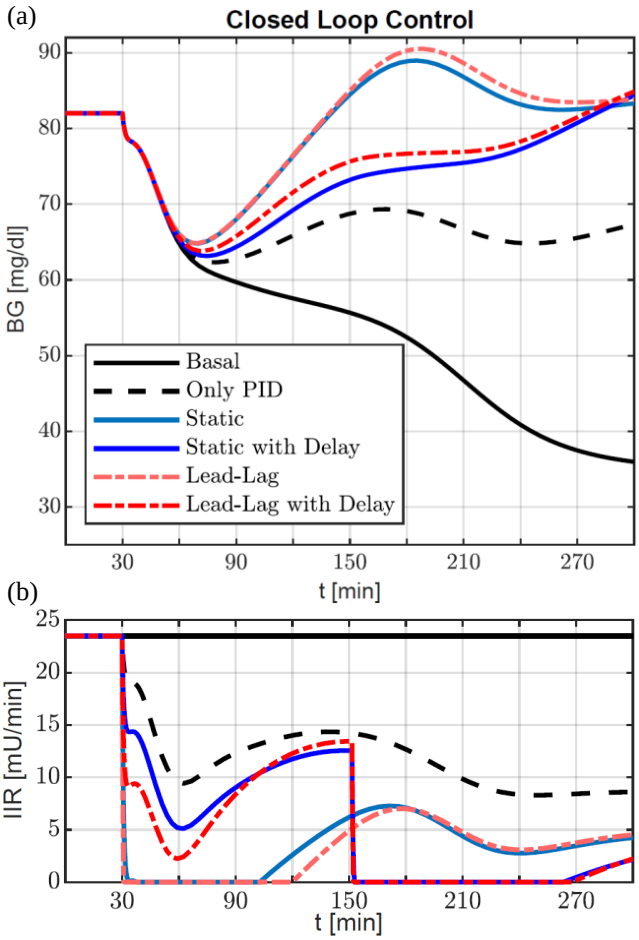


Fig. 4. (a) BG and (b) IIR during PA in closed loop control using continuation of the basal IIR, exclusive PID control and a PID controller combined with one of the developed FFCs (static, static with delay, lead-lag, lead-lag with delay).

Fig. 3(b) and Fig. 4(b) show IIR in open and closed loop control. In open loop control, FFCs without delay cause a single IIR step at the onset of PA, whereas FFCs with delay cause one initial step at the onset of PA and one delayed step approximately 2 hours later. Adding a PID controller decreases IIR during the whole PA with FFCs with delay and during the first 1.5 hours with FFCs without delay compared to open loop control. Exclusive PID control leads to a smaller reduction of IIR than PID and FFC together.

4. Discussion

The results show the potential of the developed FFCs to improve BG control during PA in both open and closed loop control by reducing time in hypoglycemia, increasing minimal BG and reducing RMSD. The initial decline in BG is a result of the delayed effect of the insulin, which was administered before the PA. As the decline occurs even with control strategies, where insulin infusion is stopped after the onset of PA, a prior announcement is necessary to prevent hypoglycemia in the examined scenario. The developed FFCs are informed about PA after the onset of PA and are thus not able to prevent the decline.

However, the FFCs are able to compensate the effect of PA after that point in time when the effect of an IIR change sets in. FFCs with delay estimate the effect of PA accurately by differing between the general effect of PA and the exhaustion effect. Hence, they compensate well the current effect of PA when the IIR change becomes effective. The FFC design does not provide means to compensate the offset, which arises because the delay of the disturbance is smaller than the delay of the process. Delays are simply neglected in that case in the FFC design. Hence, FFCs with delay do not increase BG after the initial decline, but only keep BG level constant and hypoglycemia persists. The rise in BG after $t = 210\,\text{min}$ can be explained by simplifications in the FFC design process. In contrast, FFCs without delay assume that exhaustion sets in at the onset of PA. The reduction of IIR is therefore stronger than needed to compensate the current effect of short-term PA. This causes a rise of BG till exhaustion sets in.

Combining FFCs with a PID controller improves the performance, as the PID controller causes a stronger initial IIR reduction, supporting the counteracting against the initial decline in BG, and works against the BG offset of FFCs with delay. The combination FFC and PID is superior to exclusive PID control. As the FFC input (i.e. PVO_2^{max}) changes faster than the PID input (i.e. CGM), FFCs enable an earlier adaption of IIR. FFCs allow for a more vigorous change of IIR, as they are particularly informed about PA. Moreover, FFCs can compensate the exhaustion effect, which causes a decline at $t = 180\,\text{min}$ with exclusive PID control and with continuation of basal IIR, by a proactive reduction of IIR.

Best BG control is achieved with FFCs without delay due to the low time in hypoglycemia, low RMSD and high minimal BG. An advantage of the static FFC is its simple structure, as this simplifies a potential adaption of the FFC to individuals, which might be necessary due to interpatient variability. However, the applicability of the developed FFCs is limited, as the determination of PVO_2^{max} is not feasible in everyday life. In the future, PVO_2^{max} should be replaced by another signal, e.g. the heart rate.

It should be noted that only one PA is investigated, variabilities are neglected and just one set of parameters for the PID controller and the FFCs is examined. Moreover, we expect that the exercise extension used in this paper underestimates the effect of PA, because it is derived from data of healthy subjects with the physiological reduction of insulin secretion [1] not being considered. As a high model quality is a prerequisite for significant results and a successful development of FFCs, future work should first focus on improving the model. When an enhanced model is developed, FFCs for PA should be further investigated.

Acknowledgements

This work was supervised by Prof. Dr.-Ing. Dr. med. Dr. h. c. (CTU Prague) Steffen Leonhardt, RWTH Aachen University, and Prof. Prof. h. c. Dr.-Ing. Birgit Glasmacher, Leibniz University Hannover.

Abbreviations

BG, blood glucose; CGM, continuous glucose monitoring; FFC, feedforward compensator, IIR, insulin infusion rate; PA, physical activity; PAMM, percentage active muscle mass; PT_1T_t, PT1 with delay; PVO_2^{max}, percentage maximum oxygen uptake; RMSD, root-mean-square deviation

References

[1] HÜRTER, P., DANNE, T. *Diabetes bei Kindern und Jugendlichen: Grundlagen – Klinik – Therapie*. Springer-Verlag, 2006.

[2] CASTLE, J. R., DEVRIES, J. H., KOVATCHEV, B. Future of automated insulin delivery systems. In *Diabetes Technology and Therapeutics*, 2017, vol. 19, no. 3, p. 67 – 72.

[3] TONOLI, C., HEYMAN, E., ROELANDS, B., BUYSE, L., CHEUNG, S. S., BERTHOIN, S., MEEUSEN, R. Effects of different types of acute and chronic (training) exercise on glycaemic control in type 1 diabetes mellitus. In *Sports Medicine*, 2012, vol. 42, no. 12, p. 1059 – 1080.

[4] BRETON, M.D. Handling exercise during closed loop control. In *Diabetes Technology and Therapeutics*, 2017, vol. 19, no. 6, p. 328 - 330.

[5] SORENSEN, J. T. *A Physiologic Model of Glucose Metabolism in Man and its Use to design and assess Improved Insulin Therapies for Diabetes*. Massachusetts Institute of Technology, dissertation, 1985.

[6] LENART, P., PARKER, R. S. Modelling exercise effects in type I diabetic patients. In *Proceedings of the 15th Triennial World Congress IFA*. Barcelona (Spain), 2002, p. 21 - 26.

[7] HERNÁNDEZ-ORDOÑEZ, M., CAMPOS-DELGADO, D.U. An extension to the compartmental model of type 1 diabetic patients to reproduce exercise periods with glycogen depletion and replenishment. In *Journal of Biomechanics*, 2008, vol. 41, no. 4, p. 744 – 752.

[8] DALLA MAN, C., RAIMONDO, D. M., RIZZA, R. A., COBELLI, C. GIM, simulation software of meal glucose-insulin model. In *Journal of Diabetes Science and Technology*, 2007, vol. 1, no. 3, p. 323 – 330.

[9] GUZMÁN, J. L., HÄGGLUND, T. Simple tuning rules for feedforward compensators. In *Journal of Process Control*, 2011, vol. 21, no. 1, p. 92 - 102

[10] ZIMMERMANN, M. *Implementierung eines komplexen Glucose-Insulin-Stoffwechselmodells in MATLAB*, RWTH Aachen University, Bachelor's thesis, 2012.

Glycemic Control in the Intensive Care Unit to Prevent Hyperglycemia in Critically Ill Patients

Carl-Friedrich BENNER[1]

[1]Medical Information Technology, RWTH Aachen University, Pauwelsstr. 20, 52074 Aachen, Germany

benner@hia.rwth-aachen.de

Abstract. *Within critically ill patients in the intensive care unit (ICU) stress related hyperglycemia is observed. Stress hyperglycemia is related to higher duration of stay within the hospital and an increased mortality. Clinical protocols suggest the administration of insulin to keep the blood glucose level in a normal range. A closed loop control approach allows an automated and improved insulin dosage. In this paper a mathematical model describing the glucose metabolism is adapted to simulate stress hyperglycema in critically ill patients. For an automated insulin administration two controllers are presented. The controllers are able to reduce the risk of hyperglycemia.*

Keywords

glucose control, critically ill patients, stress hyperglycemia, gycemic variablitity, closed-loop.

1. Introduction

Patients in the intensive care unit (ICU) often face stress hyperglycemia along with high glycemic variability. Stress hyperglycemia can appear after an acute illness or surgery and is characterized by blood glucose levels (BG) exceeding $140\ [mg/dl]$. The gylcemic variability describes the standard deviation of the mean BG. Both are associated with higher morbidity and mortality [1]. It can affect diabetic and non-diabetic patients [2].

Stress hyperglycemia is caused by a series of stress hormones which the human body releases after trauma, surgeries or serious illness. These stress hormones (i.e. cortisol) influence the effect of insulin on the regulation of the blood glucose level. Thereby the change of insulin sensitivity plays an important role. However, the detailed mechanisms are still unknown.

A normal BG level $80 - 120\ [mg/dl]$ within critically ill patients can be achieved by external injection of insulin. There exists no global consensus for a protocol to control BG. Different protocols were analysed in various studies targeting different ranges of BG [1][3].

In this paper an automated controller for the injection of insulin is presented. The controller should keep the BG in a specified range and reduce glycemic variability. To validate the performance of the controller a model of the human glucose metabolism is adapted to simulate hyperglycemia within critically ill patients.

2. Methods

2.1. Model of Glucose Metabolism

The model of the glucose metabolism in critical ill patients is based on the Lunze model [4]. The model describes the dynamics of BG which is controlled by the antagonistic hormones insulin and glucagon (figure 1). The specified model was modelled and parameterized for diabetic Goettingen Minipigs. Therefore, its parameters were adapted from the Dalla Man Model [5][6] according to [7]. To account for the natural blood glucose regulation the model was extended by the pancreatic insulin release which was adapted from the Sorensen model [8].

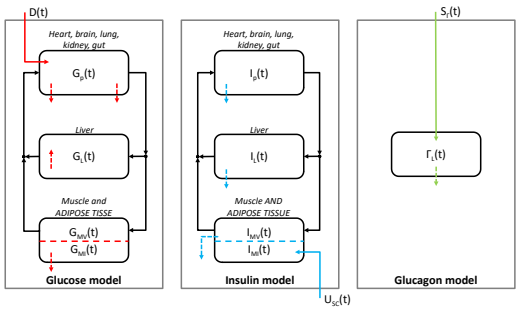


Fig. 1. Compartment model as described by Lunze [4]. The model consists of subsystems for insulin, glucose and glucagon dynamics.

Hyperglycemia was modelled as described in an extended Hovorka model [9]. A time-dependent dimensionless insulin modifier, $S_{I,MOD}(t)$, is defined, which scales the insulin sensitivity and the basal insulin concentration in the blood and the liver. Thereby the fine tuned balance of glucose, insulin and glucagon is affected allowing to simulate the BG dynamics of critically ill patients.

2.2. Feedback Controller

A feedback controller is desgined according to Aström and Hägglund [10]. The controller gets as input the intravenous measured blood glucose level. The BG level is controlled by intravenous injection of insulin. The performance of a PI controller and a PID controller are compared. According to the design rules the following parameters are derived for the PI controller:

$$k_p = 0.7; \quad T_I = 32.5\ [min] \tag{1}$$

Similarly the PID controller is characterized by:

$$k_p = 3.2; \quad T_I = 37.1\ [min]; \quad T_D = 9.3\ [min] \tag{2}$$

3. Results

In the simulation scenario several events of hyperglycemia are presented. In the first hyperglycemic phase the insulin sensitivity and basal insulin concentration are reduced to 75% compared to their normal states. In the second hyperglycemic phase the insulin sensitivity and basal insulin concentration are reduced to 50% in comparison to their normal states. In figure 2 the BG of the untreated open loop system is compared to the performance of the PI controller and PID controller.

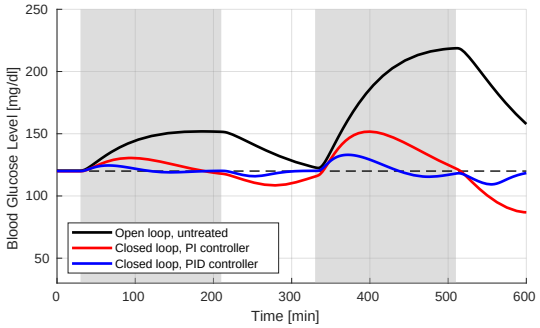


Fig. 2. During two phases (grey) the insulin sensitivity and basal insulin concentration are reduced. As a result the open loop blood glucose increases (black) while the closed loop blood glucose level remains near the target 120 $[mg/dl]$ (dashed black line) for both controllers. The performance of a PI (red line) and a PID controller (blue line) are compared.

Without external administered insulin the pancreas is not able to produce enough insulin to stabilize the blood glucose levels within the normal range. The blood glucose level rises up to 220 $[mg/dl]$. The blood glucose level normalizes after setting the insulin sensitivity and basal insulin concentration back to normal. The closed loop system with PI and PID controller is able to keep the blood glucose level within a predefined range during hyperglycemic events. The PID controller shows a marginal better performance than the PI controller. Also the PID controller shows less overshoot in blood glucose which reduces the risk of hypoglycemia, a blood glucose level below 80 $[mg/dl]$.

4. Discussion

The adapted model can simulate hyperglycemia as occuring in critically ill patients. The time response corresponds to data measured in real patients [9]. The model allows to design and to test the performance of controllers. The designed controller keeps the deviation of blood glucose level within a specified range and thereby prevents hyperglycemia and reduces glycemic variability. During the design of the control parameters a tradeoff between performance and robustness is necessary to prevent the possibilities of hypoglycemia that can occur.

Future works will focus on more complex models to describe the effect of stress hyperglycemia. Additionally, these models could be trained with patient data. Also other control concepts could be designed and evaluated with these models.

Abbreviations

BG, blood glucose; ICU, intensive care unit;

Acknowledgements

The research described in this paper was supervised by Prof. Dr.-Ing. Dr. med. Dr. h.c. Steffen Leonhardt.

References

[1] VAN DEN BERGHE, G., WILMER, A., et al. Intensive insulin therapy in the medical ICU. *NThenew england journalofmedicine*, 2006, vol. 354, no. 5, p. 449-461.

[2] PREISER, J.-C.,ICHAI, C.,ORBAN, J.-C., GROENEVELD, A. Metabolic response to the stress of critial illness. *British Journal ofAnaesthesia*, 2014, vol. 113, no. 6, p. 945-954.

[3] The NICE-SUGAR Study Investigators. Intensive versus Conventional Glucose Control in Critically Ill Patients. *New england journalofmedicine*, 2009, vol. 360, no.13, p. 1283-1297.

[4] LUNZE, K. Blood Glucose Control in Diabetic Gttingen Minipigs. *Shaker Verlag*, 2013

[5] DALLA MAN, C., RIZZA, R. A., COBELLI, C. Meal Simulation Model of the Glucose-Insulin System. *IEEE Transactions on biomedical engineering*, 2007, vol. 54, no. 10, p. 17401749.

[6] DALLA MAN, C., RAIMONDO, D. M., RIZZA, R. A., COBELLI, C. GIM, simulation software of meal glucose-insulin model. In Journal of Diabetes Science and Technology, 2007, vol. 1, no. 3, p. 323 330.

[7] ORTMANN, L. Blood Glucose Control of Type 1 Diabetes Mellitus Patients with Gaussian Process-Based Model Predictive Control. *RWTH Aachen University, master's thesis*, 2017

[8] SORENSEN, J. A Physiologic Model of Glucose Metabolism in Man and its Use to design and assess Improved Insulin Therapies for Diabetes, *Massachusetts Institute of Technology, dissertation*, 1985

[9] HOVORKA, R.; CHASSIN, L., ELLMERER, M., PLANK, J., WILINSKA, M. A simulation model of glucose regulation in the critcally. *Physiological Measurement*, 2008, vol. 29, No. 8.

[10] ASTRÖM, K. J., HÄGGLUND, T. PID Controllers: Theory, Design and Tuning. *Instrument Society of America*, 1992.

Towards Subjective Evaluation of Individualized Head-Related Transfer Functions

Shaima'a DOMA, Hark BRAREN

Teaching and Research Area of Medical Acoustics, Institute of Technical Acoustics, RWTH Aachen University, Kopernikusstr. 5, 52074 Aachen, Germany

shaimaa.doma@akustik.rwth-aachen.de, hark.braren@akustik.rwth-aachen.de

Abstract. *In augmented reality applications, individual head-related transfer functions (HRTFs) allow for realistic spatial auditory reproduction. They describe the direction-specific filtering of sound, resulting from its interaction with the listener's individual pinna, head and torso geometry, which exhibits high inter-personal variation. Different approaches for HRTF individualization aim to predict a suitable HRTF set for a user, when an individual measurement is not available for binaural synthesis. Besides an objective assessment of the resulting HRTFs, validation of these methods requires a subjective evaluation of perceptual aspects. Notably, little is known about differences in performance between the various algorithms. Therefore, a real-time HRTF demonstration tool was developed at the Institute of Technical Acoustics, RWTH Aachen University. A graphical user interface (GUI) allows the user to explore three HRTF individualization algorithms and to assess the perceptual impact of input parameters. Design and implementation aspects are presented.*

Keywords

HRTF individualization, subjective evaluation, graphical user interface, virtual acoustics.

1. Introduction

For subjective quality evaluation in spatial audio processing, conventional listening experiments focus on specific perceptual aspects, such as localization accuracy, lateralization or immersion. To have a reasonable test duration and yet be able to draw significant statistical conclusions, tested conditions have to be limited to a few. This drawback led to the implementation of a demonstration tool, which offers far more flexibility in testing three different algorithms for individualized HRTFs. While the tool is not an environment for blind listening tests, the user may investigate their own just noticeable differences (JNDs) for provided variables, the minimum audible angles (MAAs) and their overall impression for various input settings. Real-time operation is possible due to a MATLAB interface to the Virtual Acoustics framework [1], which is responsible for rendering. Upon finding a most suitable HRTF set, it may be exported at $5^\circ \times 5^\circ$ resolution. In the following, some capabilities of the HRTF demonstration tool are presented.

2. Source configuration

The concept behind the source settings panel and the output area of the GUI shows resemblance to the experiment design of Hwang et al. [2]. Given a verbal description of a source trajectory, participants were asked to assess the difference to the perceived path of a moving virtual source. In the demo tool, the user can choose between static sources and a number of predefined paths at different elevation and azimuth angles. A 3-D plot visualizes the current and previous positions of virtual sources.

3. HRTF Individualization

The idea of individualized HRTFs circumvents the necessity of acoustical HRTF measurements, e.g. when the required hardware is not available. It is interesting to evaluate differences between different approaches for individualization, as well as to measured HRTFs. The demo tool integrates three methods, for which an implementation is provided in the ITA-Toolbox [3]. The tool provides two operating modes, which are described in the following.

3.1. ITD Adjustment and Frequency Scaling

The first mode (see figure 1) makes use of an arbitrary generic HRTF, which is modified as follows:

The difference in times of arrival between the two ears, known as the interaural time difference (ITD), corresponds to a phase difference in the HRTF and is essential for azimuthal localization. The method in use (after Bomhardt and Fels [4]) models the head as an ellipsoid, with semi-major axes for head width, depth and height. A suitable ITD phase is estimated, based on an analytical solution for plane

wave incidence onto an ellipsoid, derived from the spherical solution. The algorithm replaces the ITD phase of a generic HRTF by the estimated value. In the demo tool, sliders enable a step-wise variation of the three head dimensions.

A further modification is based on an approach proposed by Middlebrooks [6]. He studies a method where spectral features of HRTFs are aligned to their predicted positions along the frequency axis, so as to fit a specific HRTF user. An expansion or compression of the frequency axis is done using a linear scaling factor. Values of a fourth slider cover the range suggested by Middlebrooks for optimal alignment of features.

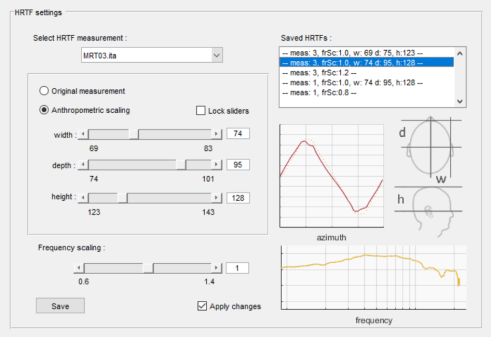


Fig. 1. HRTF settings panel for ITD adjustment and frequency scaling mode.

3.2. Reconstruction based on PCA

The second mode approximates HRTF magnitude spectra by a linear combination of so-called principal components (PCs). This approach after Bomhardt et al. [5] requires a database of HRTFs, along with anthropometric data of its members. Feeding the magnitude spectra into a Principal Component Analysis (PCA) yields a set of PCs and associated "ideal" weights for the weighted sum. For database-external users, PC weights are estimated using Multiple Regression Analysis (MRA) on six anthropometric dimensions. The demo tool user may directly compare measured and (both types of) reconstructed HRTFs for all database members (see figure 2). Besides the variation of body dimensions, sliders enable a change in the number of PCs used for reconstruction, leading to different degrees of data loss in the HRTF. The database may further be appended by additional members during operation.

4. Conclusion

An HRTF demonstration tool was presented, that provides access to three HRTF individualization algorithms. The user is familiarized with the degrees of freedom provided by the input parameters of the algorithms, and can develop an idea of their perceptual influence. Besides the direct comparison of HRTFs, a suitable set may be exported for future application.

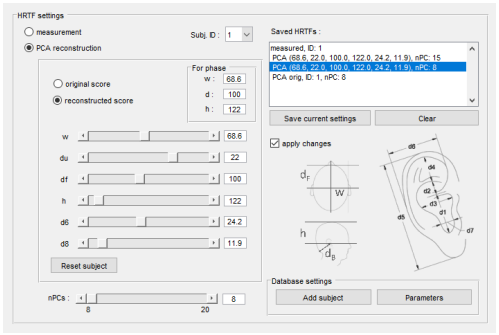


Fig. 2. HRTF settings panel for PCA mode.

Acknowledgements

Research described in the paper was supervised by Prof. Dr. Janina Fels, RWTH Aachen University, Germany.

References

[1] Institute of Technical Acoustics, RWTH Aachen University. *Virtual Acoustics - A realtime auralization framework for scientific research*, http://www.virtualacoustics.org/, accessed on 2018-09-29.

[2] HWANG, S., PARK, Y., PARK, Y.-S. Customization of spatially continuous head-related impulse responses in the median plane. In *Acta Acustica united with Acustica 96*, 2010.

[3] BERZBORN, M., BOMHARDT, R., KLEIN, J., RICHTER, J.-G., VORLÄNDER, M. The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic Measurements and Signal Processing. In Fortschritte der Akustik – DAGA. Kiel (Germany), 2017.

[4] BOMHARDT, R., FELS, J. Analytical interaural time difference model for the individualization of arbitrary Head-Related Impulse Responses. In *Audio Engineering Society Convention 137*, 2014.

[5] BOMHARDT, R., BRAREN, H., FELS, J. Individualization of head-related transfer functions using principal component analysis and anthropometric dimensions. In *Proceedings of Meetings on Acoustics 172 ASA*, Vol. 29, 2016.

[6] MIDDLEBROOKS, J. C. Individual differences in external-ear transfer functions reduced by scaling in frequency. In *The Journal of the Acoustical Society of America 106.3*, 1999.

About Authors...

Shaima'a DOMA is research assistant in the Teaching and Research Area of Medical Acoustics, RWTH Aachen University. She received her Master's degree in Electrical Engineering in 2018.

Hark BRAREN is research assistant in the Teaching and Research Area of Medical Acoustics, RWTH Aachen University. He obtained his Master's degree in Electrical Engineering in 2016, and received the Viktor and Mirka Pollak-Prize for Biomedical Engineering for his Master thesis.

Subject Movement During the Measurement of Head-Related Transfer Functions

Saskia Wepner[1], *Jan-Gerrit Richter*[1]

[1]Teaching and Research Area of Medical Acoustics, Institute of Technical Acoustics, RWTH Aachen University, Kopernikusstr. 5, 52074 Aachen, Germany

saskia.wepner@akustik.rwth-aachen.de, jan.richter@akustik.rwth-aachen.de

Abstract. *In recent years, the need for individually measured head-related transfer functions (HRTFs) has increased steadily. As measurement inaccuracies due to subject movement may influence the quality of the acquired HRTFs, best possible conditions are desirable as well as a detailed knowledge of uncertainties occurring during the measurement. In this study, it is explored how and to what extent participants move during the measurement procedure. All measurements were conducted with a fast measurement system developed at the Institute of Technical Acoustics, RWTH Aachen University. Four different types of measurements with either standing or sitting participants are compared in order to examine which type results in least movement. In one of these types, a visual real-time feedback is used to give the participants the opportunity to correct their current orientation.*

Keywords

Binaural technology, HRTF measurement, subject movement, visual real-time feedback

1. Head-Related Transfer Functions

A head-related transfer function (HRTF) describes the direction-dependent influence of the human head, torso and pinna on sound that reaches a person's eardrum. An HRTF is highly individual which results in audible influences. Therefore, it is indispensable to consider these influences in VR applications and psychoacoustic experiments that shall provide realistic scenes.

2. Measurement Set-up

HRTF measurements in this study were conducted with the HRTF measurement arc that has been developed and evaluated at the Institute of Technical Acoustics in recent years [1]. The arc is a circular array of 64 loudspeakers that cover about 160 ° of elevation (see Figure 1). It can be rotated horizontally around the person to be measured. All 16 measurements were conducted in the hemi-anechoic chamber of the Institute of Technical Acoustics, using interleaved exponential sweeps [1]. Measurements were done using the ITA-Toolbox [2].

Fig. 1. HRTF measurement arc with a participant in its centre. The arc is a circular array, consisting of 64 loudspeakers.

3. Measurement Types

Four continuous measurements [3] are compared (see Table 1). Measurement type `default` describes a set-up where the arc is rotated horizontally around the standing person without any visual feedback. For measurement type `chair`, only the person's position changes from standing to seated. For measurement type `turntable`, the person is standing on a horizontally rotated turntable. Inspired by a study by DENK ET AL. [4], a Graphical User Interface (GUI) which allows the participant to correct their current orientation, is used for the standing, unrotated participant of measurement type `display` (see Figure 2).

MEASUREMENT TYPE	POSITION	ROTATION	FEEDBACK
`turntable`	standing	person	no
`chair`	seated	arc	no
`default`	standing	arc	no
`display`	standing	arc	yes

Tab. 1. Overview on compared measurement types.

A training was carried out to get familiar with the GUI. The grey cross-hair depicts the participant's current position and orientation. Smiley faces indicate the severity of deviation from the reference position that is represented by the green cross-hair. A tolerance range of $\pm 0.8\,^\circ$ has been chosen as such deviation is not expected to result in audible differences [5]. `pitch` and `yaw` deviation is illustrated by a small blue circle that represents the person's nose. Arrows show in which direction a correction should be performed.

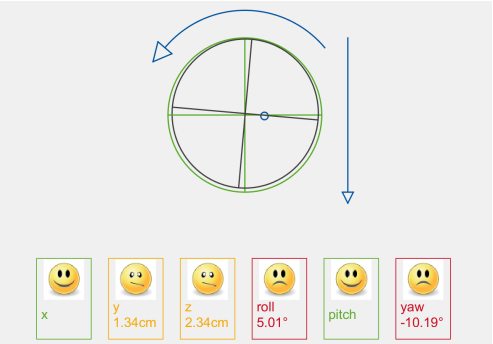


Fig. 2. Graphical User Interface for the real-time feedback.

All measurement types were carried out in a mixed order to avoid any influence resulting from the order. However, for reducing time consumption for the participant caused by reassembly of the chair or turntable, measurement type `chair` was always either the first or the last measurement. It is expected that a rotated person (`turntable`) is more strongly encouraged to move than a stationary one. Further, it is assumed that a seated person (`chair`) has less freedom in movement than a standing one and that the feedback system in `display` should help to prevent movement.

4. Results

In the following, only the results for `pitch` movement are shown in Figure 3, as `roll` and `yaw` show similar behaviour for the four measurement types. With the exception of `display`, all measurement types show a similar pattern: With increasing measurement time, participants have difficulties in keeping their original orientation.

5. Conclusion

In the scope of this study, it was found that deviations from a desired position are negligibly small. Most deviation appeared concerning orientation. The study revealed that a visual real-time feedback helps participants to reduce arbitrary movement. Additionally, it could be shown that a rotated person tends to move more that a stationary one. These findings suggest that an unrotated participant, given the possibility to correct for orientation during the measurement, leads to least inaccuracies in the resulting HRTF data.

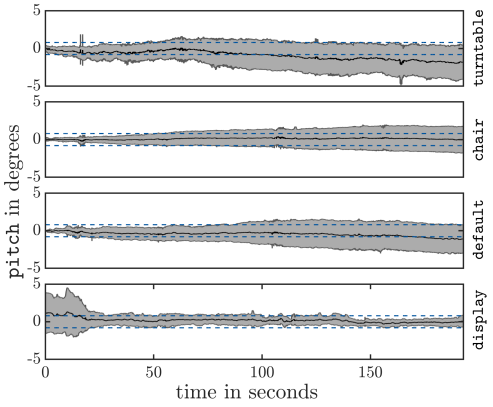


Fig. 3. Average `pitch` movement of all participants in degrees over measurement time in seconds. The blue dashed lines show the tolerances of $\pm 0.8\,^\circ$ that were accepted for deviation of the target orientation.

Acknowledgements

Research described in the paper was supervised by Univ.-Prof. Dr.-Ing. Janina Fels, Teaching and Research Area of Medical Acoustics, Institute of Technical Acoustics, RWTH Aachen University.

References

[1] RICHTER, J.-G., BEHLER, G., FELS, J. Evaluation of a Fast HRTF Measurement System. In *AES 140th Convention*. Paris (France), 2016.

[2] BERZBORN, M., BOMHARDT, R., KLEIN, J., RICHTER, J.-G., VORLÄNDER, M. The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic Measurements and Signal Processing. In *Fortschritte der Akustik – DAGA*. Kiel (Germany), 2017.

[3] RICHTER, J.-G., FELS, J. On the Influence of Continuous Subject Rotation During High Resolution Head-Related Transfer Function Measurements. In *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2019.

[4] DENK, F., HEEREN, J., EWERT, S.D., KOLLMEIER, B., ERNST, S. Controlling the Head Position during individual HRTF Measurements and its Effect on Accuracy. In *Fortschritte der Akustik – DAGA*. Kiel (Germany), 2017.

[5] BLAUERT, J. *Spatial Hearing: The Psychophysics of Human Sound Localization.* 2nd Rev. ed.: MIT Press Ltd., 1997.

About Authors...

Saskia WEPNER received her Master's degree in Electrical Engineering at RWTH Aachen University, Germany. For her Master's thesis she received the GfTA Studienpreis 2018. Since February 2019 she is part of the Medical Acoustics Group at the Institute of Technical Acoustics, RWTH Aachen University.

Jan-Gerrit RICHTER received his diploma degree in Electrical Engineering at RWTH Aachen University, Germany. Since 2013 is part of the Medical Acoustics Group at the Institute of Technical Acoustics, RWTH Aachen University.

Signal recording of 3D neurospheres on high-resolution CMOS MEA platform

Steffen Künzinger[1]

[1] BioMEMS Lab, University of Applied Sciences, Würzburger Str. 45, Aschaffenburg, Germany
stkuenzinger@gmail.com

Abstract. *The combination of high resolution signal recording of neuronal cells and the usage of human-based three-dimensional cells aggregates, allows to create a neuronal model system which is closer to the human brain functionality and its characteristics than ever before. In this study it was possible for the first time to cultivate and record signals from three-dimensional (3D) neuronal spheroids using a High Density Microelectrode array (HDMEA) chip with 4096 recording sites. It paves the way for upcoming drug tests and as a long-term goal to reduce animal testing in the pharmaceutical industry.*

Keywords

3D culture, High density microelectrode array, Neurosphere

1. Introduction

The usage of Microelectrode array (MEA) chips is a widespread method for an in-vitro investigation of neuronal network communication [1]. They are used to measure spontaneous and stimulated activity. While MEAs usually contain a maximum of 64 electrodes, the HDMEA chip contains thousands of electrodes. As a result, the density of the electrodes is much higher, which leads to a higher resolution of the electrophysiological picture of the neuronal network and therefore to a higher statistical significance, if it comes to the interpretation of the network functionality. Previous studies have already proven HDMEA chips being suitable to detect neuronal signals of retinal ganglion cells [2], dissociated neurons from cortex and hippocampi [3], and from acute murine cortico-hippocampal brain slices [4]. However, long-term studies of neuronal 3D cultures on HDMEA chips have not been described so far.

In nature, cells are embedded in a 3D environment. Various studies have shown that two dimensional (2D) neuronal cell cultures do not recapitulate structure and physiology of 3D in vivo tissues correctly [5]. In the present study we used human embryonic stem cell (hESC) derived 3D neurospheres (NS) as an model system. Previous studies already demonstrated that hESC-derived NS were electrophysiologically active using standard MEA chips [6].

2. Material and methods

2.1 HDMEA-system

The HDMEA-system, used in this study, was developed by 3Brain (Swiss). For the experiments, the HDMEA-chip "Prime" (generation: 2018) were used. On a 2,67 mm × 2,67 mm array, these chips contain 4096 electrodes, arranged in a 64 × 64 grid. The dimensions of these electrodes are 21 µm × 21 µm with a 42 µm pitch. These chips are based on CMOS and Active Pixel Sensor – technology, with a sampling rate of up to 18 kHz [7].

2.2 Neurospheres

This is a brief overview of how neurospheres are culitivated (for more information see [6]). WA09 (H9) hESC were differentiated into neural stem cells (NSC). At day 7 of neural induction primitive NSC were seeded onto ultra-low attachment 6 well plates. In contrast to the procedure described in [6], Neurobasal Plus medium was used in this study. After 18 and 36 days in suspension, NS were placed on HDMEA chips A microscopic picture of a NS after 6 days after the seeding is presented in Fig. 1.

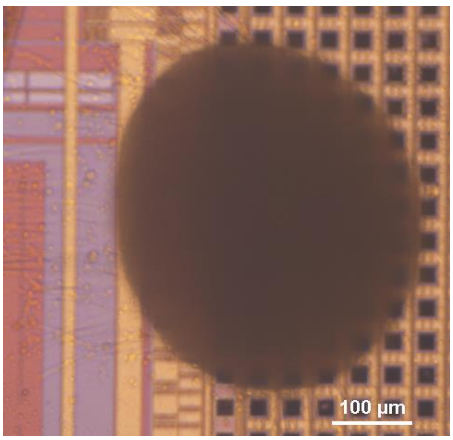


Fig. 1. ***Neurosphere on a CMOS-HDMEA***. *This 3D-Neurosphere is located at the edge of the electrode array on a HDMEA chip. The electrodes are the black squares. On the left side of the neurosphere, growing neurites and neurons are visible.*

2.3 Chip preparation

The chips were gently cleaned using a soft brush with a detergent Terg-A-zyme and sterilized properly with 70% ethanol for 20 minutes and then rinsed 4 times with double deionized water (DDW). Afterwards they were coated with Polyethyleneimine (PEI, Sigma-Aldrich, St. Louis, USA). Briefly, HDMEA chips were filled with PEI solution (0.1% in boric acid buffer) and incubated at room temperature for one hour. Then, the solution was aspirated, chips were rinsed 4 times with sterile water and air-dried overnight. At the next day, HDMEA chips were filled with Neurobasal Plus medium with supplements. The older neurospheres have been seeded on four HDMEAs and the younger once on three HDMEAs. The number of neurospheres variated from 3 to 14 NS per chip. The first measurements started after 4 DIV and took place twice a week in combination with a full medium exchange. The recorded files had a length of one minute with a sampling rate of 18 kHz.

3. Results

The first step was to check the adhesion of the NS using a microscope. A good indication for good adhesion of neurons is the presence of outgrowing neurites. At the chips with NS days which were in suspension for 36 days, the first neurites were visible on the chips 4 days after seeding. On the chips with NS which were in suspension for 18 days, the first neurites could be recognized after 14 days. This was to be expected as the older ones are already more differentiated. All in all, the coating protocol proves to be reliable.

The measurements also showed that the older NS are more mature. They showed neuronal activity after 6 days, while the younger NS did not show clear signals until after 14 days. In both cases mainly spikes and several bursts were detected. A typical recording is shown in Fig. 2. In the course of cultivation, the number of bursts increased while the signal amplitudes varied strongly with amplitudes form ± 100 µV to ± 900 µV

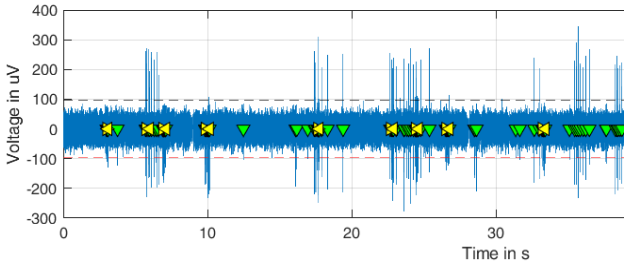


Fig. 2. ***Spike Train of neurospheral activity (section)***. *The green triangles represent a spike, while the yellow ones indicate bursts. The noise level is around ± 50 µV while the signal amplitudes are from ± 200-300 µV in this case.*

In Fig. 3 six NS are located on the HDMEA chip. By superimposing with the activity image it becomes visible that clear neuronal signals can be measured at two NS. A big, flattened NS in the middle and a smaller one at the edge are showing strong, frequently signals within a larger area of electrodes. Most of the time signals and bursts happened randomly but sometimes especially at the NS in the middle, spikes or burst happened simultaneously over the entire NS.

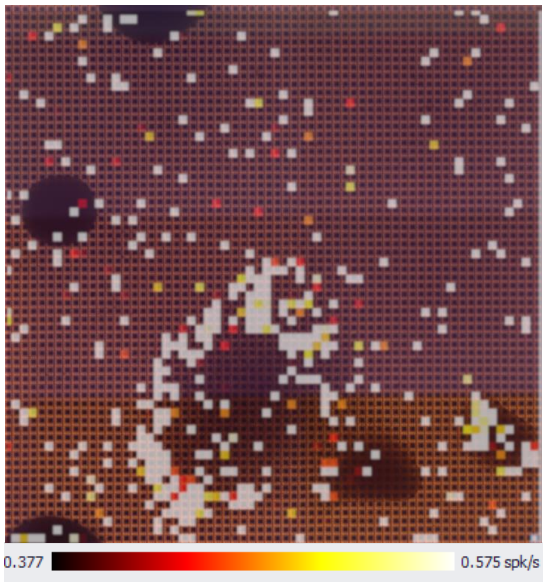


Fig. 3. ***Neuronal activity over a HDMEA***. *Two images (microscope and amplitude image) are superimposed here Each pixel represents an electrode of the HDMEA. There is a high activity (yellow/white) in the large oval area on the bottom half and the smaller area near the bottom-right corner. Neurospheres were present exactly in these areas.*

The number of active electrodes per NS depends on the size of the NS. In this experiment, the average was 80 active electrodes per NS. For comparison, the MEA have an average of 17 active electrodes (see [6]).

4. Conclusion and Outlook

A 3D culture like neurospheres can be cultivated on HDMEA chips for several weeks allowing analysis of network development .

Despite a high noise level of ± 50 µV, signals are well recognizable due to the good signal-to-noise ratio of mostly 3 – 8. MEA chips that are used in the BioMEMS Lab usually have a signal-to noise ratio from 2 – 4.

Spikes can be positive or negative. This means that it is recommended to look at both directions to detect spikes for the signal.

HDMEAs allow neural signals to be derived from more electrodes per neurosphere compared to MEA, which leads to a higher statistical significance.

To record all 4096 channels at 18 kHz generates a huge amount of data. A one-minute recording creates a file that is 8,4 GB large. In order to analyze these data, high demands must be placed on hardware and software.

This gathered knowledge will later be applied to future experiments like drug screening applications or the understanding of connectivity.

Acknowledgements

Author thanks Prof. Dr.-Ing. C. Thielemann, BioMEMS Lab, University of Applied Sciences Aschaffenburg, who supervised the research described in the paper. Furthermore, he thanks Dr. rer. nat. M. Mayer for her help with cell handling, the advice on signal evaluation and fruitful notes.

This project was funded by the Federal Ministry of Education and Research in the frame of the project iNeuTox (13FH516IX6).

References

[1] PINE, J. Recording action potentials from cultured neurons with extracellular microcircuit electrodes, *J. Neurosci. Methods*, 1980, vol. 2, no. 1, p. 19–31.

[2] MACCIONE, A., GANDOLFO, M., ZORDAN, S., Microelectronics, bioinformatics and neurocomputation for massiveneuronal recordings in brain circuits with large scale multielectrodearray probes, *Brain Research Bulletin*, 2015, vol. 199, part B, p. 118 - 126

[3] BERDONDINI, L., IMFELD, K., MACCIONE, A. Active pixel sensor array for high spatio-temporal resolution electrophysiological recordings from single cell to large scale neuronal networks, *Lab on a Chip*, 2009, vol. 9, no. 18, p. 2644 - 2651

[4] FERREA, E., MACCIONE, A., MEDRIHAN, L. Large-scale, high-resolution electrophysiological imaging of field potentials in brain slices with microelectronic multielectrode arrays, *Frontiers in Neural Circuits*, 2012, vol. 6, article. 80, p. 1 - 14

[5] HUANG, G., LI, F., ZHAO, X. Functional and Biomimetic Materials for Engineering of the Three-Dimensional Cell Microenvironment, *Chem. Rev.*, 2017, vol 117, p. 12764-12850

[6] MAYER, M., ARRIZABALAGA, O., LIEB, F. Electrophysiological investigation of human embryonic stem cell derived neurospheres using a novel spike detection algorithm, *Biosensors and Bioelectronics*, 2018, vol. 100, p. 462-468

[7] IMFELD, K., NEUKOM, S., MACCIONE, A. Large-Scale, High-Resolution Data Acquisition System for Extracellular Recording of Electrophysiological Activity, *IEEE transactions on biomedical engineering*, 2008, vol. 55, no. 8, p. 2064 – 2073.

About Author

Steffen KÜNZINGER was born in Aschaffenburg (Germany). He graduated with a bachelor´s degree in industrial engineering at the university of applied sciences in Aschaffenburg. Currently he studies for a master´s degree, where he faces research questions in the field of cell signal derivation in the BioMEMS Lab.

Functionalized Hydrogel as Glucose Imaging Sensor for *in vitro* Cell Cultures

Jens Jüttner[1], *Lukas Kress*[1]

[1] BioMEMS Lab, Faculty of Engineering, University of Applied Sciences, Würzburger Str. 45, Aschaffenburg, Germany

s140415@th-ab.de, s140473@th-ab.de, biomems@th-ab.de

Abstract. *In this proof-of-principle study, we present an innovative fluorescence-based glucose sensor system that allows the continuous measurement of glucose in cell cultures. The enzyme-free glucose sensor chemistry, which can be stored for long periods, can easily be integrated into standardized labware, e.g. well plates. Possible application could be long term monitoring of cell metabolism.*

Keywords

In vitro cell systems, optical sensor, glucose, PDMS, bioprinting, fluorescence microscopy, ConA, Dextran

1. Introduction

The field of tissue engineering aims to create 2D and 3D artificial cellular tissue for regenerative medicine or for organ on a chip models. A major challenge is the sufficient supply with nutrients for long-term cultivation. Therefore, enough nutrients such as glucose must be in medium. To detect the amount of glucose in media, a novel enzyme-free sensor chemistry, which was initially intended for continuous glucose measurement for diabetes patients, was developed [1].

This alternative approach of detecting glucose in cell culture media is the competitive complex formation with Concanavalin (conA) and Dextran. The receptor ConA and the competitor Dextran are labelled with fluorescent dyes, which, in the case of joint complex formation, exhibit resonance-based quenching of their fluorescence (Fluorescence resonance energy transfer, FRET). The displacement of Dextran by glucose leads to an increased distance between the two fluorophores (donor dye on Dextran and acceptor dye on conA), which prevents the non-radiative energy transfer of the FRET-effect. Therefore, the dyed Dextran is enabled to emit fluorescence again. Thus, the intensities of the dye markers function as indicators for the glucose concentration. By reducing the glucose concentration in media, the effect described above can be reversed. This can be measured with appropriate fluorescence photometers, e.g. integrated into a microscope. Furthermore, a third dye can be used as a marker to provide a reference point during measurements [2, 3].

2. Materials and Methods

2.1 Hydrogel Preparation

Fluorescently labeled ConA (Zedira GmbH, Darmstadt, Germany) and Dextran (Life Technologies GmbH, Darmstadt, Germany) were loaded to previously formed alginate microbeads by incubation of the beads in a solution of 4 mg/ml ConA and 7 mg/ml Dextran. The reference dye was embedded into polystyrene microbeads. The exact nature of the fluorescent dyes used cannot be disclosed because of a patent in preparation.

Alginate beads containing the sensor chemistry were mixed with the polymer Nelfilcon™ (kindly provided by CIBA Vision GmbH, Grosswallstadt, Germany), an ultraviolet-polymerizable, derivatized poly(vinyl alcohol). The highly viscous solution was molded into poly(methyl methacrylate) molds and was cured under ultraviolet light. After that, the cured material was separated into small, < 1 mm thin discs with a diameter of 4 mm, so called WOMIs (German; Enhanced Ocular Mini Implants).

2.2 Standard Labware Integration

As shown in Figure 1, the WOMIs were embedded in bioprinted PDMS cages on a 6-well plate to avoid floating. These cages were fabricated with a 3DDiscovery™ bioprinter (regenHU, Villaz-Saint-Pierre, Switzerland). The PDMS cage consisting of 30% w/v Sylgard 184 (DOWSIL, Midland, USA) and 70% w/v SE 1700 (DOWSIL, Midland, USA) was printed with a 26 G dosing needle, 9 mm/s feed rate and 300 kPa pressure. Afterwards, PDMS was cured for 48h at room temperature to gain final form stability. Curing of the PDMS at higher temperatures was not possible due to the sensor chemistry not being able to withstand > 50 °C.

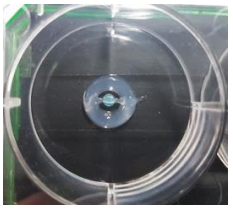


Fig. 1. A single well of a 6-well plate into which a WOMI embedded into a PDMS cage was integrated. The PDMS strand securing the WOMI into its position can be seen in the center.

2.3 Glucose Imaging Setup

WOMIs integrated into a 6-well plate were exposed to DMEM cell culture medium containing different concentrations of glucose (0-10 g/l in steps of 1 g/l, 15 g/l and 20 g/l) for 30 min each. For every concentration, samples were examined using a standard fluorescence microscope with fluorescence color filters for the emission of the donor dye (bandpass: 620 nm) and for the emission of the reference dye (bandpass: 750 nm). The excitation wavelength emitted by a mercury lamp was restricted to 590 nm due to an excitation filter (bandpass: 590 nm).

2.4 Image Analysis and Value Formation

A custom MATLAB-tool was used for image analysis. Firstly, fluorescence calibration images of the 6-well plate were subtracted from 16-bit greyscale images of each color filter (donor and reference). Afterwards, the quotient donor/reference of both summed up pixel intensities was determined. This value is an indicator for the concentration of glucose in media.

3. Results

Fluorescence images as exemplarily shown in Figure 2a and 2b were captured to generate a sensor response curve using the MATLAB-tool described in the previous chapter.

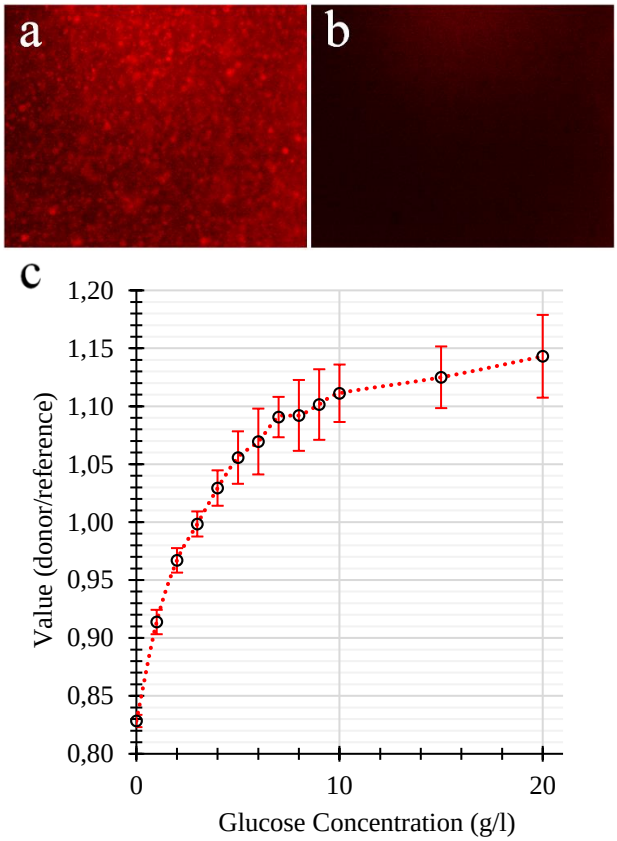


Fig. 2. Fluorescence-response of a WOMI. a) Donor at 10 g/l b) Reference signal at 10 g/l c) Obtained calibrated sensor curve (n=6): The values represent the mean values and the dashed line shows an interpolation between the data points.

As can be seen in the first half of the curve shown in Figure 2c, the slope from 0 g/l to 5 g/l is significant and thus shows high sensitivity of the biosensor in this region. Once it reaches 5 g/l, the curve starts to flatten out and sensitivity decreases.

4. Discussion

In this proof-of-principle study, we have introduced an optical glucose sensor applicable to *in vitro* cell cultures with possible integration into standard labware. It enables simple and noninvasive read out of glucose concentrations up to 10-20 g/l in media with a standard fluorescence microscope. This can reduce the required amount of medium, as it can be determined when new cell culture medium is needed due to low glucose levels.

Furthermore, as shown in Figure 2c, the fluorescence-response curve fits very well to its initial purpose of blood sugar monitoring in the human body (typ. 0.5 – 1,2 g/l) [4] as the sensor chemistry has a high sensitivity in this range.

In contrast to enzyme-based glucose sensors, a chemical equilibrium and no reaction rate is used to determine the glucose concentration, which renders these ConA/Dextran based sensors more robust against cross sensitivities. In addition, the system offers an excellent shelf life of more than 1 year when stored dry and ensures an *in vitro* application of more than 4 weeks.

5. Conclusion and Outlook

In this work, we have successfully developed a simple method to determine glucose concentrations within an *in vitro* cell culture medium using a standard fluorescence microscope. Furthermore, the system provides the advantages described in chapter 1 and 4, e.g. reversibility or longevity.

In the future, the PDMS cage could be made redundant by functionalizing the surface of the well plate with double bonds, e.g. by plasma treatment, enabling the direct print of Sensorsol, the basic material the WOMIs are made of, in standard labware [5]. Since the sensor targets cell cultures, it may be useful to combine it with a hydrogel-based O_2 sensor chemistry to determine not only the glucose level but also the O_2 content of the culture medium [6].

Acknowledgements

The authors thank Prof. Dr.-Ing. C. Thielemann, BioMEMS Lab, UAS Aschaffenburg, who supervised the research described in the project. This work was funded by our partner EyeSense GmbH, Grossostheim, Germany and project GlukZ AZ-1253-16 of the Bavarian Research Foundation. Furthermore, EyeSense GmbH provided us with the sensor chemistry and fruitful advice. We also thank M. Eng. S. Allig for conducting preliminary experiments.

References

[1] MUELLER, A. J., KNUTH, M., NIKOLAUS, K. S., KRIVANEK, R., KUESTER, F., HASSLACHER, C., AUFFARTH, G. U. Blood Glucose Self-Monitoring with a Long-Term Subconjunctival Glucose Sensor. *Journal of Diabetes Science and Technology*, jan 2013, 7(1):24-34.

[2] BALLERSTADT, R., SCHULTZ, J. S. A fluorescence affinity hollow fiber sensor for continuous transdermal glucose monitoring. *Analytical chemistry*, sep 2000, 72(17):4185-4192.

[3] MUELLER, A. J., KNUTH, M., NIKOLAUS, K. S., HERBRECHTSMEIER, P. First Clinical Evaluation of a New Long-Term Subconjunctival Glucose Sensor. *Journal of Diabetes Science and Technology*, jul 2012, 6(4):875-883.

[4] HANSEL, R., STEINEGGER, E., STICHER, O. Pharmakognosie – Phytopharmazie. 6. Ausgabe, apr 2013, ISBN: 9783662092699.

[5] LAI, J., SUNDERLAND, B., XUE, J., YAN, S., ZHAO, W., FOLKARD, M., MICHAEL, B. D., WANG, Y. Study on hydrophilicity of polymer surfaces improved by plasma treatment. *Applied Surface Science*, mar 2006, 252(10):3375-3379

[6] TRAMPE, E., KOREN, K., AKKINENI, A. R., SENWITZ, C., KRUJATZ, F., LODE, A., GELINSKY, M., KUEHL, M. Functionalized Bioink with Optical Sensor Nanoparticles for O_2 Imaging in 3D-Bioprinted Constructs. *Advanced functional materials*, sep 2018, 28(45).

About Authors

Jens JÜTTNER was born in Aschaffenburg (Germany), where he graduated with a bachelor´s degree in Industrial Engineering at the University of Applied Sciences Aschaffenburg. Currently he is studying for a master´s degree as a part of the Master Research Programme. His current research projects in the BioMEMS-Lab focus on the fields of bioprinting and 3D-Cell Culture.

Lukas KRESS was born in Fulda (Germany). He graduated as Bachelor of Engineering at the university of applied sciences Aschaffenburg. Additionally, he extended his bachelor studies by a Double Degree Programme with the University of Applied Sciences Seinäjoki, Finland. Currently, he is participating in the Master Research Programme at UAS Aschaffenburg, where he faces research questions in the field of bioprinting and brain on a chip sensor systems in the BioMEMS Lab.

ICA decomposition as a basis for finding hidden EEG structures by dimension reduction methods

Natalie BROZOVA[1], *Marek PIORECKY*[1, 2]

[1] Dept. of Biomedical Technology, Czech Technical University in Prague, nám. Sítná 3105,Kladno, Czech Republic
[2] National Institute of Mental Health, Topolová 748, Klecany, Czech Republic

natalie.brozova@fbmi.cvut.cz, marek.piorecky@fbmi.cvut.cz

Abstract. *Electroencephalography (EEG) is the electrical signal of the brain whose analysis makes it possible to recognize its pathologies and brain diseases. EEG signal is very complex, hence for better understanding several linear and non-linear processing techniques are widely used to help to detect the hidden data structure. The aim of this study is to test linear and nonlinear dimensional reduction to obtain information about the internal hidden data structure. Based on this information, suitable properties for describing the EEG signal can be determined. Our method seems to be working out well for alpha activity and eye activity space examination.*

Keywords

EEG, dimension reduction, ICA, PCA, tSNE, DBSCAN.

1. Introduction

Dimension reduction techniques are used to preserve the dimensions of the data which contains only the information important for further analysis [1]. With EEG signal, dimension reduction techniques are mostly used on the whole EEG record or on the segment's features and then the result is classified [2, 3]. Many applications also include blind source separation, which is nowadays widely performed by independent component analysis (ICA). The result of ICA can also be used as input data for dimension reduction [3, 4]. In this study, the result of ICA is described by five commonly used features, normally used for EEG signal description. This feature space is then being reduced. At the same time, a nonlinear method of reduction to raw data is used, which attempts to detect artifact residues in records. This could help to determine appropriate features for EEG signal description and suppressing unwanted artifacts, which would help the doctor to not be burdened by the feature selection.

2. Methods

EEG records are analyzed in two groups – 6 epileptic patient and 10 controls. The study protocol has been approved by the Hospital Na Bulovce ethical committee. The EEG data are 19-channel-ambulatory EEG recordings measured in 10-20 system EEG on the Brain-Quick digital system. MATLAB 2015b were used for all signal processing.

2.1 Independent component analysis

The EEG signal is segmented on three minutes long segments, on which ICA is performed. ICA contains extraction and separation of statistically independent sources of a biological signal, therefore it helps to reveal hidden inner data structure [5]. Each independent component should represent some specific EEG activity or noise. It also preserves topographical information, therefore it is a good technique for determining eye artifacts [6]. In this study, fast ICA algorithm is used. It is using simple estimations of negentropy based on a principle of maximal entropy. Negentropy can be understood as a measure of "non-normality" (non-gaussian distribution). This algorithm is based on mutual information, which can be written as:

$$I(s) = \int f_s(s) log \frac{f_s(s)}{\prod f_{s_i(s_i)}} ds, \quad (1)$$

where f_s are probability densities. The result of ICA is obtained by minimalizing mutual information [7].

2.2 Dimension reduction and classification

Independent components are then described by five features from time and frequency domain. Following features were selected: autocorrelation, focal topography, focal trial activity, dominant frequency, and intensity weighted mean frequency (IWMF). Principal component analysis (PCA) was applied to the criterial – component space. PCA is a linear dimension reduction technique, which projects data in reduced space, defined by orthogonal principal components, which are linear combinations of original variables. PCA also sorts principal components by the size of the total variance explained by each component [8]. Coefficients of

variables to determine principal components are stored in the Loading matrix. The loadings are calculated by singular value decomposition (SVD) of covariance matrix C [8]:

$$C = \frac{X^T X}{I-1} = LS^2L^T = Z\Lambda Z^T, \quad (2)$$

where $Z\ (J \times J)$ is an orthogonal matrix, $S\ (J \times J)$ is a diagonal matrix with nonzero singular values on its diagonal, $L\ (J \times J)$ is the loading matrix, which collects on each *j*-th column the coefficients of the J variables for defining the *j*-th principal component. $X\ (I \times \mathrm{J})$ is an original matrix with I samples and J variables and Λ is the diagonal matrix which contains the nonnegative eigenvalues of decreasing magnitude, each eigenvalue encodes the variation related to the corresponding component [8].

As the nonlinear method was used t-SNE which was applied on criterial-component space and was also used for topographical map space reduction. This method counts the probability of spatial distribution in high dimensional and low dimensional space, then it tries to optimize these two probabilities [9]:

$$p_{i|j} = \frac{e^{-d(x_i,x_j)^2/2\sigma_i^2}}{\sum_{k\neq i} e^{-d(x_i,x_k)^2/2\sigma_k^2}}. \quad (3)$$

Finally, by the density-based classification method, DBSCAN (see Algorithm 1), in both reduced spaces, it is evaluated if it is possible to separate data into clusters, describing any specific EEG activity or artifact [10]. Information from linear and nonlinear reduction is compared.

Data:
Clusters collected from nodes - CC,
distance - ε,
minimum number of points to create dense region - *minPts*

```
begin
    for each cluster C in CC do
        for each point P in C do
            if P is visited then
                Continue to next P
            end
            else
                mark P as visited
                nbrPts ⟵ points in ε-neighborhood of P that are not in C
                if sizeof(nbrPts) ≥ minPts then
                    Merge C with every cluster, to which points nbrPts belongs
                end
            end
        end
    end
end
```

Algorithm 1. Pseudocode demonstrates process of the DBSCAN algorithm in simpler way [11].

3. Results and discussion

3.1 PCA

Topographical maps in Fig. 1 correspond to component space created by ICA. Criteria from the frequency domain have higher variance, so we assume they are better for separating hidden layers. DBSCAN defined 4 clusters in 2D PCA space for controls and 3 clusters for the epileptic group, see Fig 2. In the control group, the first cluster corresponded to mixed activity mainly with low dominant frequency, the second cluster represented alpha activity, third beta activity, and eye movement activity was clearly separated in the fourth cluster. The difference between the control group and the group of epileptic patients was found in the spectrum on 10 and 15 Hz. Components of the epileptic group have some dominant frequencies at 10 Hz and higher frequencies but mostly lower than controls. In the epileptic group, the 2D space was not as visibly separated as the control group 2D space, hence DBSCAN could not separate all the clusters. It formed the first large cluster of mixed activity. The second and the third clusters correspond to eye movement activity. Clusters corresponding to eye movement activity both in controls and the epileptic group can be characterized by high autocorrelation. The percentages of total variance explained by each principal component are also evaluated – the first component of each group explains the majority of information of the original data.

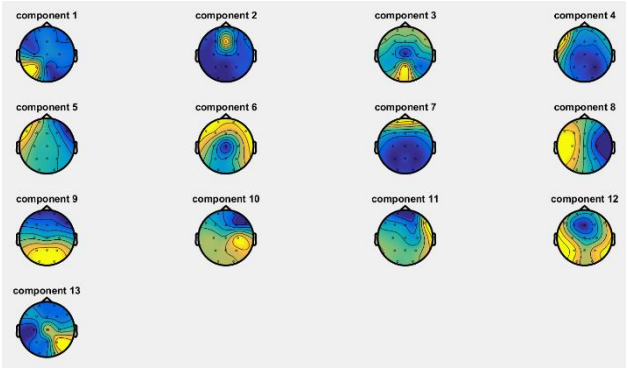


Fig. 1. Topographical maps of independent components – control group.

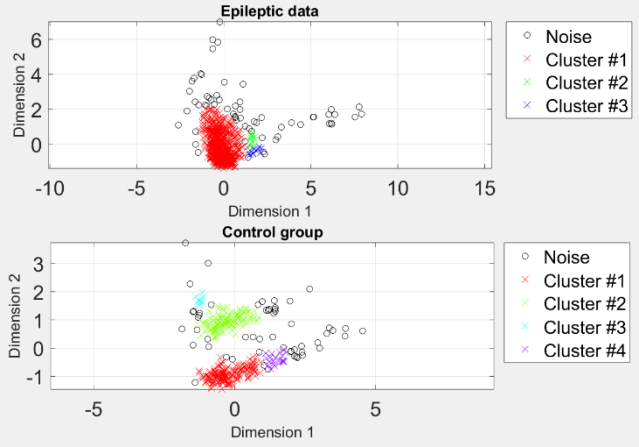


Fig. 2. 2D spaces reduced by PCA, classified by DBSCAN.

3.2 tSNE

After reducing the high dimensional space to 2D space using tSNE, DBSCAN defined 3 clusters for controls and 5 clusters for the epileptic group (Fig 3). tSNE created clusters based rather on intensity weighted mean frequency (IWMF) than dominant frequency. The clusters in the control group

don't define any specific activity but are clearly separated by IWMF. In the epileptic group the first and second clusters corresponded to mixed activity, second with higher IWMF. The third cluster clearly separated eye movement activity with high autocorrelation, fourth cluster epileptic components with high dominant frequencies and IWMF. Fifth mixed cluster corresponded to epileptic components and other components with high dominant frequencies and IWMF.

For describing eye movement artifacts autocorrelation proved to be relevant, see Fig 4. and Fig 5.

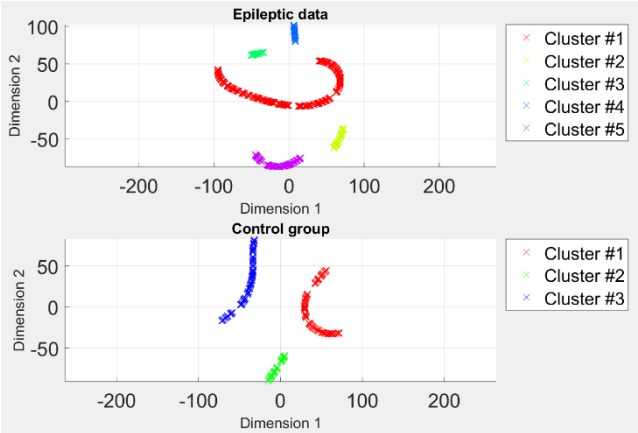


Fig. 3. 2D space reduced by tSNE, classified by DBSCAN

Topographical map space after t-SNE was diverse to 16 – 20 clusters based on DBSCAN parameters. Eye artifacts were also dominant here. It is also possible to consider the use of Riemannian geometry as an optimal concept of EEG space. In future work, linear Euclidean metrics (used by current algorithms) could be compared with curved Riemannian space.

4. Conclusion

This study seeks to test linear and nonlinear dimension reduction techniques on the feature space of independent components of the EEG signal. A nonlinear method was also used for evaluation distribution of topographical map (representing EEG amplitude distribution), without specific feature assumption. Two-dimensional space was classified by density-based method DBSCAN because space consists of nested clusters. Given features used on ICA components and reduced by PCA in control group describe separately alpha activity, beta activity. In both control and epileptic group, the features describe eye movement artifacts. When reduced by tSNE, given features in epileptic group describe the epileptic activity and eye movement artifacts.

We found out that PCA describe more precisely controls space and t-SNE seems beater for epileptic patients' dataset. In the future, it would be essential to check the impact of epileptic graphoelements on nonlinear character of EEG space.

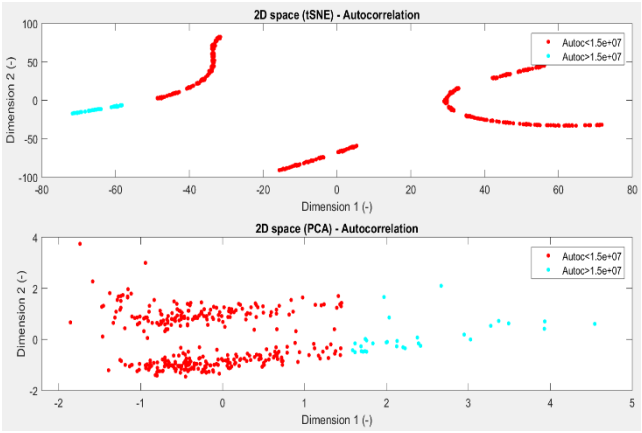


Fig. 4. 2D spaces reduced by tSNE and PCA, described by autocorrelation – control group.

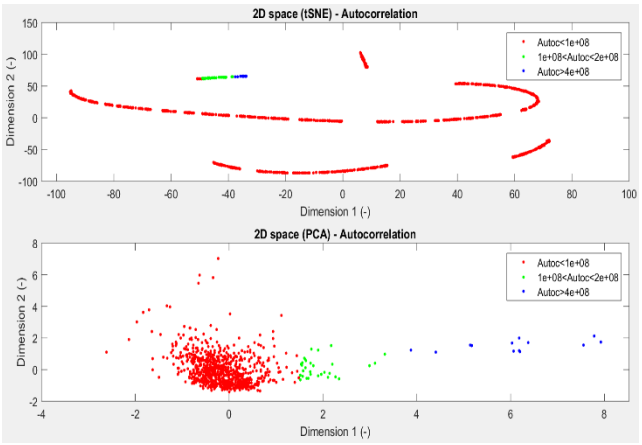


Fig. 5. 2D spaces reduced by tSNE and PCA, described by autocorrelation – epileptic group.

Features from frequency domain – dominant frequency and intensity weighted mean frequency proved to be relevant as well for describing various physiological EEG activities and epileptic activity.

Acknowledgements

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant no.SGS18/159/OHK4/2T/17. We thank to MUDr. Svojmil Petranek and Bulovka Hospital, Dep. of Neurology, Prague, Czech Republic.

References

[1] CARREIRA-PERPINAN, M. A Review of Dimension Reduction Techniques. Sheffield, 1997. *Technical Report. Dept. of Computer Science University of Sheffield*.

[2] SO-YOUN, P., JU-JANG, L. EEG Feature Extraction and Classification Using Data Dimension Reduction. *IEEE International*

Conference on Industrial Informatics. 2008, 6, 355-358. DOI: 10.1109/INDIN.2008.4618123.

[3] ARTONI, F., DELORME, A., MAKEIG, S. Applying dimension reduction to EEG data by Principal Component Analysis reduces the quality of its subsequent Independent Component decomposition. *NeuroImage*. 2018, 175, 176-187. DOI: 10.1016/j.neuroimage.2018.03.016.

[4] JAMES, Ch., HESSE, Ch. Independent component analysis for biomedical signals. *Physiological Measurement*. 2004, 26(1), 15-39.

[5] LEE, Te-W., GIROLAMI, M., SEJNOWSKI, T. Independent Component Analysis Using an Extended Infomax Algorithm for Mixed Subgaussian and Supergaussian Sources. *Neural Computation*. 1999, 11(2), 417-441.

[6] CHAUMON, M., BISHOP, D., BUSCH, D. A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of Neuroscience Methods*. 2015, 250(7), 47-63. ISSN 0165-0270.

[7] NAIK, Ganesh a Dinesh KUMAR. An Overview of Independent Component Analysis and Its Applications. *Informatica*. 2011, 35, 63-81.

[8] HOTELLING, H. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*. 1933, 24(6), 417.

[9] BIRJANDTALAB, J., BARAN POUYAN, M., NOURANI, M. Nonlinear Dimension Reduction for EEG-Based Epileptic Seizure Detection. 2016 *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI).* 2016, 595-598. DOI: 10.1109/BHI.2016.7455968.

[10] KHAN, K., REHMAN, S., AZIZ, K., FONG, S., SARASVADY, S. DBSCAN: Past, Present and Future. *The Fifth International Conference on the Applications of Digital Information and Web Technologies*. 2014, 232-238.

[11] MERK, A., CAL, P., WOZNIAK, M. Distributed DBSCAN Algorithm – Concept and Experimental Evaluation. *Proceedings of the 10th International Conference on Computer Recognition Systems CORES 2017. CORES 2017. Advances in Intelligent Systems and Computing, vol 578. Springer*. DOI: 10.1007/978-3-319-59162-9_49

About Authors...

Natalie BROZOVA is studying bachelor degree and finishing her final thesis on Reducing the dimension of EEG space as the basis for datamining.

Marek PIORECKY is currently studying doctoral studies with a focus on integration of EEG and fMRI signals through symmetric approach and dimension reduction.

Airwave oscillometry

Tomáš VLČEK

Dept. of Biomedical Technology, Faculty of Biomedical Engineering, Czech Technical University, Nám. Sítná 3105, 272 01 Kladno, Czech Republic

tomas.vlcek@fbmi.cvut.cz

Abstract. *Airwave oscillometry (AOS) is a new non-invasive method for lung function testing. In this work, I compared AOS with spirometry in experimental measurements on a group of 20 volunteers divided into asthmatics and healthy individuals. The R5 parameter, measured by AOS, was correlated with the FEV1 parameter, measured by spirometry. The results show a strong correlation between the parameters for healthy subjects (r = 0.85) and a moderate correlation for the asthmatics (*r *= −0.50). In the asthmatic group, R5 was increased and FEV1 decreased when compared with the healthy group and the reference data. AOS has the potential to recognize respiratory system disease and serve as a substitute for spirometry.*

Keywords

Airwave oscillometry, spirometry, forced oscillation technique, airway resistance, *FEV1*, asthma

1. Introduction

Spirometry is the standard and elementary method of functional lung examination. The method is based on the principle of measuring the volume flow in time. Spirometry examination is characterized by forced breathing maneuvers such as quick inspiration and expiration. The maneuvers require the cooperation of patients and their coordination with instructions of clinical staff [1, 2]. Airwave oscillometry (AOS) was designed to prevent forced maneuvers. The AOS method is based on the Forced Oscillation Technique (FOT), where fine acoustic waves with pressure amplitude of about 1–2 cmH2O propagate through the respiratory system at nine different frequencies (5–37 Hz) [3]. Acoustic waves at lower frequencies from 5 to 17 Hz advance deeper into the central airways, where they are reflected and returned to the device. Conversely, frequencies from 19 to 37 Hz reach the peripheral airways only. The examination requires only resting breathing, which reduces a patient load and is important when examining, for example, small children or seniors who would not be able to perform required spirometry maneuvers [4, 5].

The aim of the study was to compare functional lung examination by spirometry and AOS in a group of healthy volunteers and in a group of volunteers with asthma.

2. Methods

The prospective intervention study was approved by the Ethics Review Committee of the Faculty of Biomedical Engineering. The study involved 20 volunteers; all participants knew the experimental protocol of the study and signed an informed consent with their participation in the study. Volunteers were chosen from among students of the Faculty of Biomedical Engineering. The participants were questioned about the present and past status of their respiratory system and about their sports activities. Based on their answers, the volunteers were divided into the asthma group or the healthy group. Basic description of the participants is summarized in Tab. 1.

Number of probands (health/asthma)	**20 (10/10)**
Men	**11**
Women	**9**
Mean Age	**21.9**
Mean Height (cm)	**177.1**
Mean Weight (kg)	**78.2**
Mean BMI (kg/m²)	**24.9**

Tab. 1. Characteristics of study participants.

The experimental protocol follows the instrument manuals supplied by the manufacturer [3, 6]. Spirometry was performed with Ergostik (Geratherm Respiratory GmbH, Germany). The AOS measurements were made with tremoFlo C-100 (THORASYS Thoracic Medical Systems Inc., Canada). The spirometry examination began with quiet inspiration and expiration, followed by rapid maximum inspirations and forced expirations. The AOS measurement consisted of resting breathing for a period of 16 seconds. Both examinations were repeated at least 3 times for each participant.

Among the measured parameters, I focused primarily on the FEV1 parameter for spirometry and on the R5 parameter for AOS, as these variables classify the function of the respiratory system. FEV1 represents the forced

expiratory volume exhaled with the maximum effort within 1 s after the maximum possible inspiration. The parameter asses the gas flow rate of exhalation and its decrease below 80% of the reference is a sign of lung disease [7]. R5 represents the airway resistance at the frequency of 5 Hz of acoustic waves that pass through the respiratory system. Increased values of R5 may indicate the impaired function of the respiratory system [3]. I averaged results from repeated measurements of individual parameters for each proband and, using the Pearson correlation coefficient (*r)*, see formula (1), I calculated the mutual dependence of these two parameters in both groups of participants. Normality of the data was verified by the Q-Q graph (quantile-quantile graph) and by the Lilliefors test of normality ($\alpha = 0.05$) using Statistics and Machine Learning Toolbox in Matlab (MathWorks, USA). The data were processed in Excel 2016 software (Microsoft Office, USA).

$$r_{X,Y} = \frac{s_{xy}}{s_x s_y} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad (1)$$

3. Results

The relationship between the FEV1 parameter of spirometry and the R5 parameter of AOS is presented in Fig. 1. Pearson correlation coefficient is $r = -0.50$ for the asthma group and $r = 0.85$ for the healthy group.

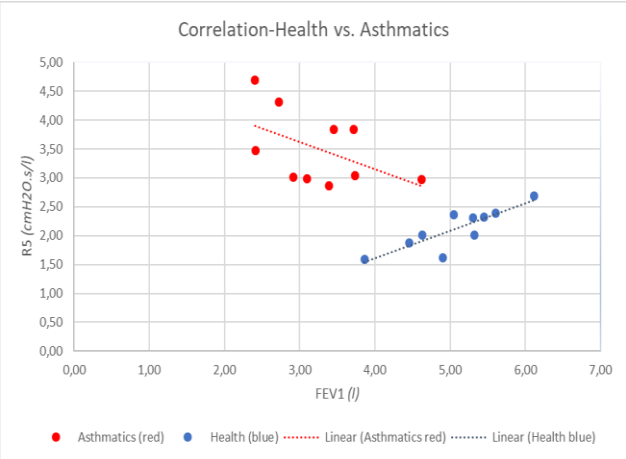


Fig. 1. Summary correlation diagram for healthy and asthmatic patients. Data points represent mean FEV1 and R5 for each volunteer.

4. Discussion and Conclusion

The work focused on the relationship between R5 and FEV1. In the asthmatic group, the correlation coefficient was negative. This confirms a respiratory disease as the expired volume decreases with an increase of the respiratory tract resistance. The respiratory tract is narrowed by the surrounding smooth muscle or clogged with mucus, which increases the airway resistance and limits the airflow [8]. On the contrary, in the group of healthy individuals, the correlation coefficient was positive. Forced expiratory effort leads to positive pressure in the thoracic cavity that can narrow the airways and thus increase airway resistance [9]. The results of the work could be more significant and accurate if the study included more participants. In asthmatics, different severity of disease in participants could contribute to the low correlation coefficient. However, the results correspond with the expected behavior of airways in healthy individuals and asthmatics. When comparing the measured values of R5 and FEV1 with the reference values found in literature, the healthy group corresponded to standard physiological values. Asthmatics, on the other hand, showed elevated R5 values and reduced FEV1 values.

Both spirometry and AOS are non-invasive methods of functional lung examination. The advantage of AOS is the simplicity of the examination process, where only the quiet breathing of a patient is required. The advantage of spirometry is a wide range of measured parameters to evaluate the pathological conditions of the respiratory system. This work shows that AOS may be a good alternative that corresponds with spirometry and has the potential to recognize a disease of the respiratory system and thus serve as a substitute for spirometry. That may be clinically relevant especially when a patient is unable to perform forced maneuvers and spirometry cannot be used.

Acknowledgements

We thank Ing. Jakub Ráfl, Ph.D., for supervising the bachelor thesis and Ing. Jan Matějka and Ing. Markéta Masopustová for their valuable advice on the thesis. We also thank the probands for their efforts and time and the Faculty of Biomedical Engineering for the measuring devices. The work was supported by grant SGS17/203/OHK4/3T/17 of the Czech Technical University in Prague.

References

[1] MARWA, A. Repeatability of respiratory impedance and bronchodilatory response in asthmatic children. [Diploma thesis]. Halifax, Nova Scotia, Canada: Dalhousie University, 2016.

[2] Metody měření průtoku plynů – Lékařské přístroje a zařízení. [online, in Czech]. 2018 [cited 10. 5. 2018]. Available from: https://sites.google.com/site/lpz2011123/zobrazovaci-metody/spirometr/metody-mereni-pruatoku-plynua.

[3] *tremoFlo C-100 airwave oscillometry system user manual*. Thoracic Medial Systems Inc., 2015.

[4] ČÁP, P., BIČÍKOVÁ, K., PAUL, T. Impulsní oscilometrie – nová metoda měření plicních funkcí. [in Czech]. *Alergie*, 2000, vol. 4, no. 2.

[5] CLEMENT, J., DUMOULIN, B., GUBBELMANS, R., HENDRIKS, S., VAN DE WOESTIJNE, K. P. Reference values of total respiratory resistance and reactance between 4 and 26 Hz in

children and adolescents aged 4–20 years. *Bull Eur Physiopathol Respir*, 1987, vol. 23, p. 441–48.

[6] *Ergostik – Blue Cherry User Manual Version 1.2.1*. Geratherm Respiratory GmbH., 2012.

[7] Spirometry: Procedure, normal values, and test results. *Healthline: Medical information and health advice you can trust*. [online]. 2005 [cited 19. 9. 2017]. Available from: http://www.healthline.com/health/spirometry#side-effects4.

[8] Astma. Wikipedie: Otevřená encyklopedie. [online, in Czech]. 2018 [cited 8. 4. 2018]. Available from: https://cs.wikipedia.org/w/index.php?title=Astma&oldid=15869358.

[9] BRAVENÝ, P., FIŠER, B., HELLER, J., HERGET, J., HONZÍKOVÁ, N. *Poznámky k přednáškám z fyziologie (vol. 1): Srdce a krevní oběh, dýchání, ledviny, acidobazická rovnováha, svaly*. 2nd ed. Jinočany: H & H, 1992.

About Author

Tomáš VLČEK is an undergraduate student of the program Biomedical technician at the Faculty of Biomedical Engineering.

Analytical Review of Erbium-Ytterbium Waveguide Amplifier for C-Band Allocated DWDM System

Tomáš HUSZANÍK[1]

[1] Dept. of Electronics and Multimedia Communications,Faculty of Electrical Engineering and Informatics,Technical University of Košice, Košice, Slovakia

tomas.huszanik@tuke.sk

Abstract. *The extensive growth of telecommunication systems had led the development of active optical networks. Multichannel DWDM (Dense Wavelength Division) systems are the most promising fiber optical technology today. The existence of long-haul DWDM systems is possible due to fiber amplifiers. In this paper the performance and gain characteristics of erbium-ytterbium co-doped waveguide amplifier (EYDWA) is presented using 980 nm pump wavelength. The densisity of ytterbium ions and laser pump power is allternated to find the optimal amplifier parameters to be suitable for multichannel DWDM transmission.*

Keywords

Erbium, Er-Yb waveguide amplifier, DWDM, Ytterbium.

1. Introduction

Over the past few years the use of dense-wavelength-division multiplexing (DWDM) technology has driven the development of broad-band and code-transparent optical devices, such as multiplexers, splitters, couplers, combiners and amplifiers. The most of current DWDM systems use the modern EDFAs made out of silica-based glass doped by Er^{3+} ions. Limits of EDFAs used in DWDMs are well known. The main drawback of EDFA is its complexity and thus its price. Erbium doped waveguide amplifier (EDWA) represents a good alternative to EDFAs. EDWAs have several advantages over commonly used EDFAs such as: no crosstalk between simultaneously amplified WDM channels, low noise level and low power consumption. Unlike EDFA, thanks to its smaller size, EDWA offers high integration options. It can be integrated with other optical components such as: splitters, couplers, combiners, etc. in one chip. However, EDWA may suffer for the small dimensions of integrated devices, because the overall gain is given by the length of the doped waveguide. To compensate this, the waveguide requires doping with high concentration of Er^{3+}. EDWA is usually pumped at 980 nm wavelength [1][2]. This pump, with the large absorption cross-section cause the absence of a spontaneous emission of the waveguide. It provides good performance of optical gain, efficiency and low noise level. Performance of EDWA can be optimized using ytterbium ions – Yb^{3+}. Amplifier of this kind is called erbium-ytterbium co-doped waveguide amplifier (EYDWA). Ytterbium absorption band overlaps with the erbium absorption band – 980 nm. Compared to erbium, ytterbium has larger absorption cross-section at 980 nm. So, it is often used to increase the sensitivity of erbium [3][4].

In this paper, a simulation model of EYDWA is presented at different configurations. The ability of amplifying multiple DWDM channels is also considered. The structure of this paper is as follows: chapter two provides the description of a simulation model and results are discussed in chapter three.

2. Simulation model

The simulation model of an erbium-ytterbium co-doped waveguide amplifier was created in OptiSystem™ environment. The simulation is based on TDSS (Time-Domain Split-Step) principle [5]. Experimental realization of an erbium-ytterbium co-doped waveguide amplifier is shown in Fig. 1. This amplifier is made to amplify a single 193.1 THz (1552.524 nm) optical signal.

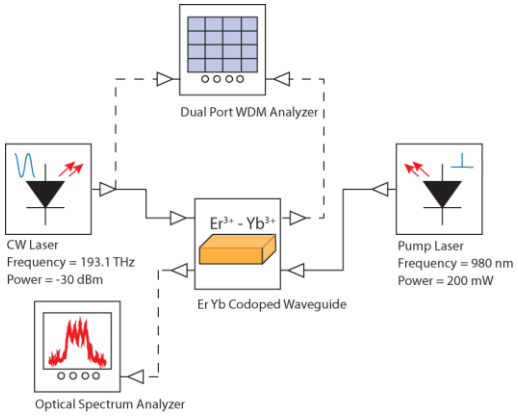


Fig. 1. Experimental setup of co-doped erbium-ytterbium waveguide amplifier with counter-directional pump.

Losses due to the introduction of Er^{3+} and Yb^{3+} ions in the material by diffusion at the signal and pump

wavelength are considered in the simulation model. The pump laser wavelength is set to 980 nm. The waveguide is pumped co-directionally. The pump power is lineary swept from 10 to 300 mW. Density of Er^{3+} ions in the waveguide is 2e+026 atoms/m^3 throughout all simulations. Density of ytterbium ions Yb^{3+} is swept from 1e+026 to 2e+027 atoms/m^3 to observe the optimal amount of ytterbium needed in the waveguide. Fig. 2 shows absorption parameter of Er^{3+} and Yb^{3+}. Ytterbium ions have higher absorption parameter which cause decrease of amplifier gain when the concentration of Yb^{3+} ions is increased. It means that the pump power will be absorbed by Yb^{3+} ions at the end of the waveguide causing decrease of amplifier's gain [6-8]. Parameters of proposed co-doped EYDWA are in Tab. 1.

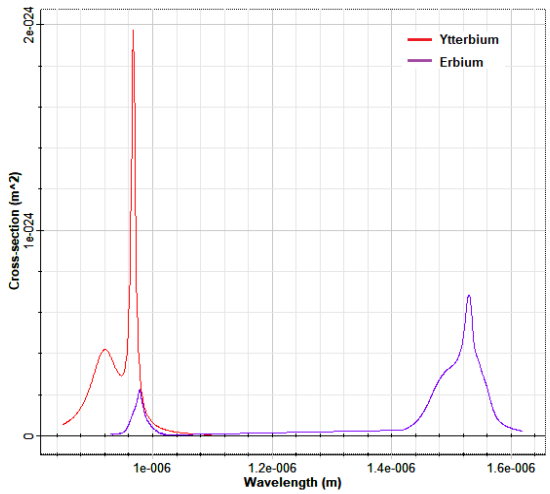


Fig. 2. Absorption parameters of erbium and ytterbium.

Parameter	Value	Unit
Er ion density	2e+026	atoms/m^3
Er metastable lifetime	11	ms
Er signal excess loss	29.3	dB/m
Er pump excess loss	37.3	dB/m
Yb iod density	1e+026	atoms/m^3
Yb metastable lifetime	29.1	ms
Yb signal excess loss	24	dB/m
Yb pump excess loss	30.5	dB/m
Waveguide length	3	cm
Pump signal wavelength	980	Nm
Pump signal power	200	mW

Tab. 1. Er-Yb waveguide amplifier parameters.

In the second simulation setup co-doped EYDWA for multichannel amplifying is presented. In this setup, 8-channel, 16-channel, 32-channel and 64-channel DWDM signal is amplified using co-doped EYDWA. The main aim of this setup is to determine efficiency of proposed co-doped EYDWA to amplify signal of a large bandwidth depnding on the density of Yb^{3+} ions. This experimental setup is shown in Fig. 3. DWDM channels of this setup are allocated in optical C-band (1530 nm – 1565 nm) which has the lowest attenuation and is compatible with fiber amplifiers. Channel frequencies of 8-channel DWDM setup are (193.10 – 193.80) THz (1552.52 – 1546.92 nm), 16-channel DWDM – (193.10 – 194.60) THz (1552.52 – 1540.56 nm), 32-channel DWDM – (193.10 – 196.20) THz (1552.52 – 1527.99 nm), 64-channel DWDM (190.10 – 196.40) THz (1577.03 – 1526.44 nm). Channel spacing between adjacent channels is set to 100 GHz for 8,16, 32 and 64 channel DWDM respectively.

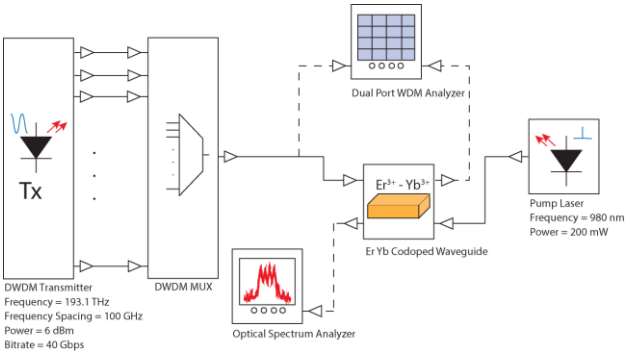


Fig. 3. Experimental setup of codoped erbium-ytterbium waveguide amplifier with counter-directional pump with DWDM transmitter.

3. Model simulation results and discussion

First, the performance of single wavelength channel amplified by EYDWA was analyzed. This experimental setup is shown in Fig. 1.

In Fig. 4, a signal amplification with a wavelength of 1552.524 nm can be seen as a function of ytterbium concentration for different lengths of the amplifier waveguide. Wavelength of 1552.524 nm (193.1 THz) is used due to it is the reference wavelength for DWDM systems. In this case, concentration of Yb^{3+} ions is swept from 1e+026 to 2e+027 atoms/m^3. Density of Er^{3+} remains the same: 2e+026 atoms/m^3. In this experiment, five lengths of active waveguide is analyzed: 1cm, 2cm, 3cm, 4cm and 5 cm. By increasing the density of Yb^{3+}, the gain of the 1 cm waveguide amplifier increases by a small portion. The gain is 0.514 dB for 1e+026 atoms/m^3 density of Yb^{3+}. The maximum gain of 0.840 dB is achieved for 2e+027 atoms/m^3 density. The gain difference is 0.326 dB. The gain of EYDWA is increasing as we increase the density of ytterbium ions for 2, 3 and 4 cm waveguide. However, as we reach the critical length of 5 cm and ytterbium density of 1.4e+027, the gain starts to decrease. The reason is that the ytterbium clusters are formed. It causes the energy to not transfer to erbium ions and energy of the pump laser will be wasted as the signal reaches the end of the waveguide. Based on these results, the 4cm waveguide provides the best gain performance. Minimum and maximum gain of each amplifier as well as the gain difference is provided in Tab. 2.

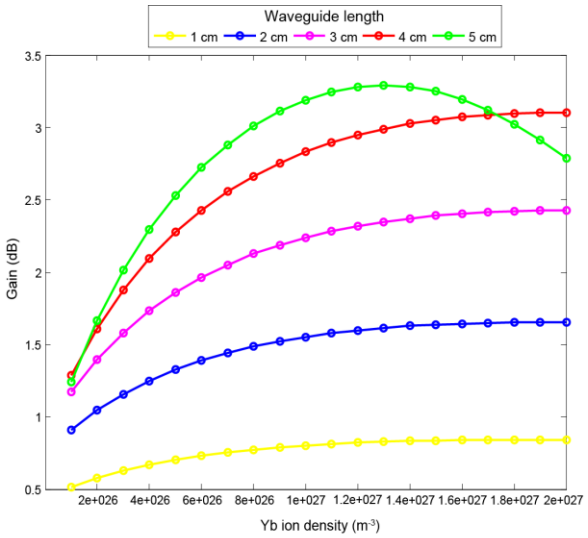


Fig. 4. Er-Yb waveguide amplifier gain figure using different Yb ion density waveguide.

EYDWA waveguide length (cm)	Min gain (dB)	Max gain (dB)	Gain difference (dB)
1	0.514	0.840	0.326
2	0.908	1.652	0.717
3	1.171	2.426	1.255
4	1.287	3.102	1.815
5	1.243	2.788	1.545

Tab. 2. Values of gain of EYDWA for experiment no.1.

Fig. 5 shows the EYDWA gain as a function of the pump power. The setup is the same as in the previous experiment, however, the Yb^{3+} ion density is set to 1e+027 atoms/m^3. The pump power is swept from 10 to 300 mW with the linear step.

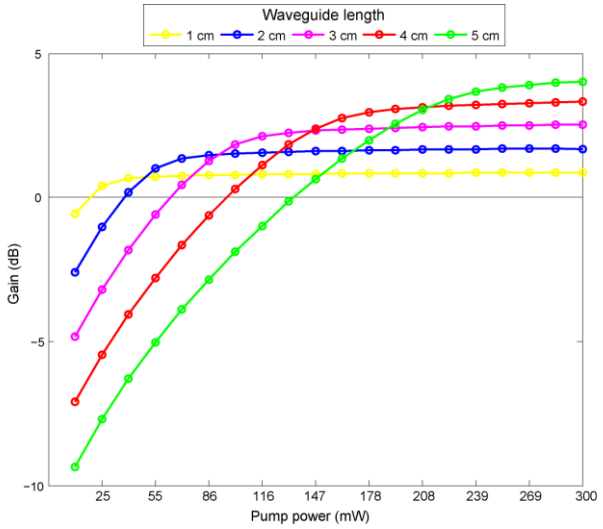


Fig. 5. Er-Yb waveguide amplifier gain figure using different pump power.

Based on these results, the EYDWA with 5 cm waveguide is the least efficient in term of the total gain at lower pump power followed by 4 cm waveguide amplifier. 1cm and 2 cm waveguide amplifiers provide relatively stable output gain for 71.05 mW and up to 300 mW pump power. Table of minimum and maximum gain is provided below (Tab. 3).

EYDWA waveguide length (cm)	Min gain (dB)	Max gain (dB)	Gain difference (dB)
1	-0.556	0.871	1.427
2	-2.597	1.716	4.313
3	-4.843	2.533	7.376
4	-7.104	3.315	10.419
5	-9.366	4.024	13.390

Tab. 3. Values of gain of EYDWA for experiment no.2.

Next, the ability of the proposed co-doped EYDWA to amplify multiple wavelength channels of DWDM is tested. For this simulation, following 40 Gbps DPSK-modulated (Differential Phase Shift Keying) DWDM systems were formed: 8, 16, 32 and 64 channels. Channel spacing is 100 GHz. Channels are allocated in the optical C-band, see Tab. 4.

DWDM channels	1st channel frequency (THz)	1st channel wavelength (nm)	Last channel frequency (THz)	Last channel wavelength (nm)
8	193.10	1552.52	193.80	1546.92
16	193.10	1552.52	194.60	1540.56
32	193.10	1552.52	196.20	1527.99
64	190.10	1577.03	196.40	1526.44

Tab. 4. DWDM channels allocation.

In this case, the EYDWA setup is following: waveguide length is 4 cm and pump power is 200 mW. Fig. 6 and Fig. 7 show combined signal and noise spectra of 16 and 32- channel DWDM signal amplified by proposed EYDWA. For both cases, noise reaches approximately -70 dBm respectively.

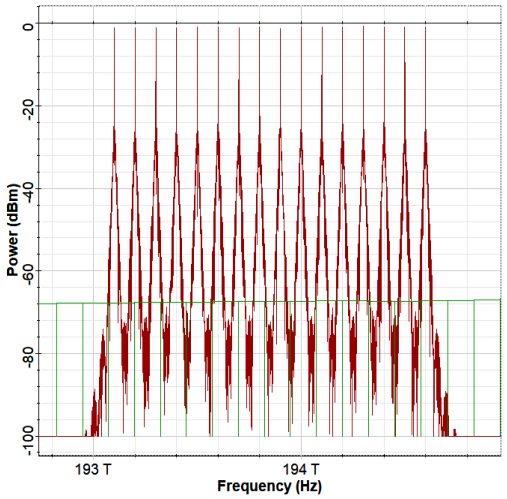


Fig. 6. Output optical spectra of 16-channel DWDM.

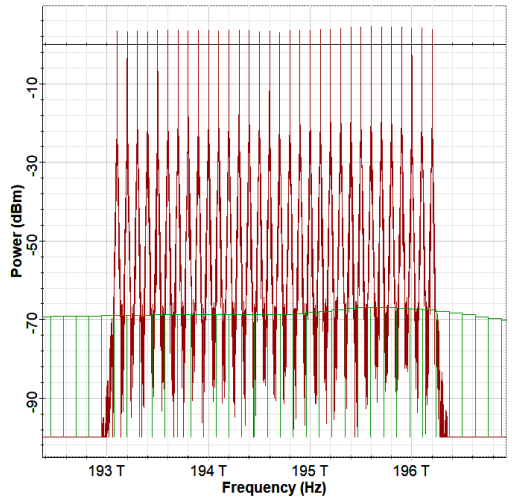


Fig. 7. Output optical spectra of 32-channel DWDM.

Fig. 8 shows the total gain dependence of Yb^{3+} ion density (swept form 1e+026 to 2e+027 atoms/m^3) for 8, 16, 32 and 64-channel DWDM. The total gain efficiency of proposed co-doped EYDWA is decreasing with increased number of channels. The bandwidth of 8-channel DWDM is 700 MHz and 16 channel 1500 MHz. EYDWA is able to amplify 8-channel and 16-channel signals effectively with increasing Yb^{3+} ion density. However, total gain of EYDWA drops when amplifying larger bandwidth signals: 32-channel (3100 MHz) and 64-channel (6300 MHz). Values of gain of this experiment are in Tab. 5.

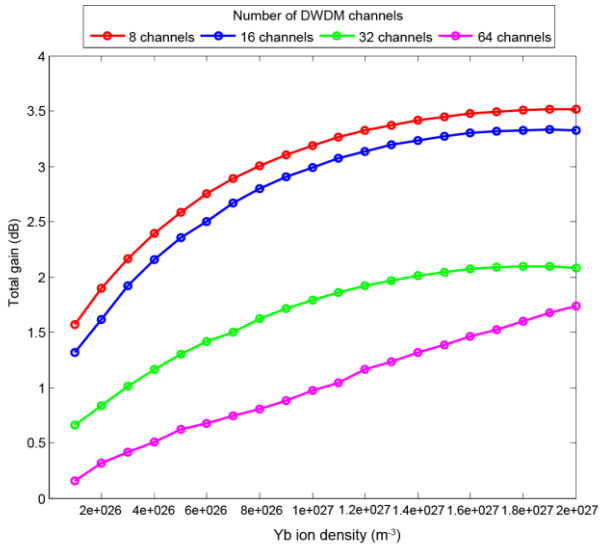


Fig. 8. Er-Yb waveguide amplifier total gain figure with different number of DWDM channels.

DWDM channels	Min gain (dB)	Max gain (dB)	Gain difference (dB)
8	1.565	3.512	1.947
16	1.313	3.323	2.01
32	0.661	2.083	1.422
64	0.154	1.736	1.582

Tab. 5. Values of gain of EYDWA for experiment no.3.

4. Conclusion

In this paper, the performance of erbium-ytterbium co-doped waveguide amplifier for DWDM system is presented. Based on the obtained results, these conclusions can be made: the output gain of EYDWA is increasing by increasing the density of ytterbium ions in the waveguide. However, the total gain depends on the length of the waveguide. The optimal length of the waveguide in these experiments is 3 cm and 4 cm, as it provides sufficient gain level. Compared to 3 and 4 cm waveguide, 1cm and 2 cm waveguide EYDWA have lower overall gain at the same Yb^{3+} ion density. With regards to the simulation results, 5 cm waveguide amplifier is at least suitable for DWDM as it is less efficient at lower pump power levels and its gain decrease with increased Yb^{3+} ion density. The performance of EYDWA amplifying multichannel signal is also analyzed. Based on this experiment, proposed EYDWA is less efficient with 32 and 64-channel DWDM deploying 100 GHz channel spacing.

Acknowledgements

Research described in this paper was supervised by Ján Turán and Ľuboš Ovseník. This work was supported by following research grants: KEGA 023TUKE-4/2017, VEGA 1/0772/17 and APVV-17-0208 - Resilient mobile networks for content delivery.

References

[1] COKRAK, A. C., ALTUNCU, A. Gain and Noise Figure Performance of Erbium Doped Fiber Amplifiers (EDFA). In *Journal of Electrical & Electronic Engineering*, 2004, vol. 4, no. 2, p. 1111 – 1122.

[2] RASHEED, B. O. Gain Improvement of Erbium –Ytterbium Waveguide Amplifier at Different Pump. In *Contemporary Engineering Sciences*, 2012, vol. 15, no. 2, p. 75-83.

[3] KAMINOW, I. P., LI, T., WILLNER, A. E. Optical Fiber Telecommunications VIB: Systems and Networks, 6th edition, 2013, p.1148, ISBN: 9780123972378.

[4] GOEL, A., MISHRA, R. S., Design of broadband erbium-doped fiber amplifier using very high-speed integrated circuit hardware description language for next-generation optical network. In *Optical Engineering*, vol. 50, no. 1, 2011, pp. 1-6.

[5] HUSZANÍK, T., TURÁN, J., OVSENÍK, L. Utilization of 10 gbps DWDM system with duobinary modulation into passive optical network. In *Journal of Communications Software and Systems*, 2018, vol. 14, no. 4, p. 367-375. DOI:10.24138/jcomss.v14i4.644.

[6] HUSZANÍK, T., TURÁN, J., OVSENÍK, L. Impact of the optical fiber nonlinear phenomenon on the 16-channel DWDM OC-768 long-haul link. In *Elektrotehniski Vestnik/Electrotechnical Review*, 2018, vol. 85, no. 5, p. 255-262. ISSN: 0013-5852.

[7] DUTTA, M. K., Design and Performance Analysis of EDFA and SOA for Optical WDM Networks: A Comparative Study. In *14th IEEE India Council International Conference (INDICON)*, Roorkee, 2017, pp. 1-6. DOI: 10.1109/INDICON.2017.8487827.

[8] IVANIGA, T., P. Ivaniga, "Comparison of the Optical Amplifiers EDFA and SOA Based on the BER and -Factor in C-Band", *Advances in Optical Technologies*, vol. 2017, 2017, ID 9053582, pp.1-9, doi.org/10.1155/2017/9053582.

Analysis of Reflector Size Impact on Modeled Helix Antenna for 5.2 GHz Usable for Experimental FSO/RF Hybrid System

Michal MÁRTON[1], *Michal ŠPES*[2], *Jakub URBANSKÝ*[3]

[1] Department of Electronics and Multimedia Communications, Technical University of Košice, Boženy Němcovej 32, 040 01 Košice, Slovakia

[2, 3] Department of Electrical Engineering, Technical University of Košice, Mäsiarska 26, 040 01 Košice, Slovakia

[1]michal.marton@tuke.sk, [2]michal.spes@tuke.sk, [3]jakub.urbansky@tuke.sk

Abstract. *We are living in the ages of innovations. We are surrounded by a lot of communication technologies. Each application needs to be plugged into network. Offline devices became useless because they can´t offer relevant data in the time. Users want to be connected to network and communicate with other users. In recent years, devices primary not intended for communication need to be gradually optimized on equipment with new technologies, enabling them to communicate between each other or across different platforms. For example we are surrounded by a lot of smart technologies (smart watches, smart glasses, smart wristbands, etc.). We need to react on users demand and their devices which need to be plugged into network. Expansion of such systems has been supported by invention of optical fiber as transmission medium. Optical fibers are capable of high data rates and long distance transmission. Non-fiber optical transmission is an idea researched and analyzed by many research activities. The FSO (Free Space Optics) system is considered as the main system representing communication system based optical atmospheric transmission. We realized experimental model of FSO system in area of Technical University of Košice. Based on knowledge and experiences about FSO we decided to design and consequently implement backup link for this system, because optical beams are sensitive to weather changes. The main point of this paper is to compare several 5.2 GHz helix antennas designs suitable for our experimental FSO/RF system.*

Keywords

FEKO, FSO/RF system, helix antenna.

1. Introduction

Communication systems are expanded through many areas. Research groups are focused to improve the possibility of communication and connectivity. There are two approaches available. The first approach is focused on improvement of availability of communication system by optimization of transmission channel. The second approach is focused on the using different method of transmission with using differently transmission channels. The FSO systems have many advantages, such as wide bandwidth ensured by optical atmospheric transmission and high level of security. Attack on the optical beams may cause the connectivity loss. FSO system operates in non-licensed frequency range without needs special permissions on the other hand, this system has also some drawbacks. The main drawback is the sensitivity to weather changes. Weather dominantly affects transmission of optical beams because the transmission channel is atmospheric with all inhomogeneities which cause the interference and scattering. The wavelength transmitted signal is small compared to water particles. We have realized FSO system in the area of Technical University of Košice. Our optical link consists of one pair of optical heads placed in line of sight. Distance between optical heads is about 400 m [1-8].

2. Models of helix antennas

The design of helix antennas is realized in special software package used for analysis of electromagnetic compatibility named FEKO. FEKO contains three subprograms. Software named CAD FEKO is used for design process. Parameters of the reference antenna are in the Tab. 1 [9-16].

Parameter	Value
Frequency	5.2 GHz
Number of turns	5
Diameter of helix	4.4 cm
Diameter of conductor of helix	0.2 cm
Diameter of reflector	6.2 cm
Length of helix	12.2 cm

Tab. 1 Parameters of the reference antenna.

For comparison, seven designs of the helix antenna are proposed using CAD FEKO. They are shown in Fig. 1.

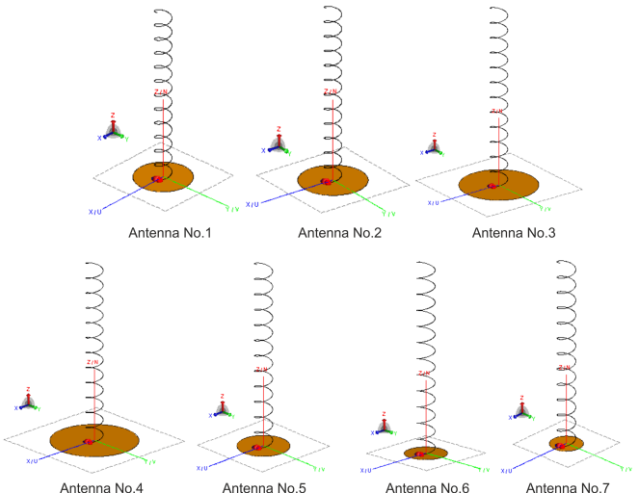


Fig. 1 Models of designed antennas.

The main difference between them is the reflector diameter. The aim of this experimental setup is to monitor the impact of this parameter on antenna´s performance. The values of diameter of helix antennas are in Tab. 2.

Variable	Diameter of reflector
Antenna No.1	6.346
Antenna No.2	7.346
Antenna No.3	8.346
Antenna No.4	9.346
Antenna No.5	5.346
Antenna No.6	4.346
Antenna No.7	3.346

Tab. 2 Different value of key parameter of the model.

3. Analysis of experimental results

The analysis of obtained results is provided by software called POST FEKO. Radiation patters of antennas are in Fig. 2.

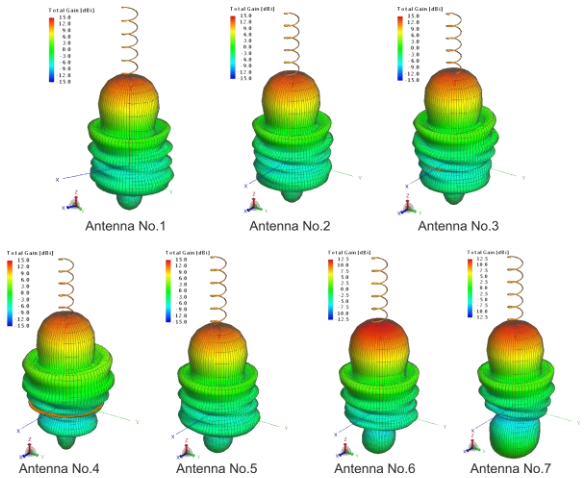


Fig. 2 Radiation pattern of designed antennas in 3D view.

The radiation pattern placed in polar coordinates is in Fig. 3 for angle $\Phi = 0°$.

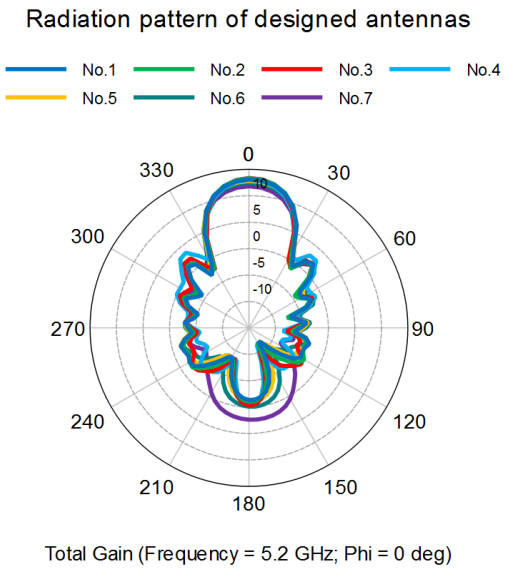


Fig. 3 Radiation of modelled antennas for $\Phi = 0°$.

Obtained results of total gain of antennas are provided in Tab. 3 In this table, maximal values of obtained gain denoted are in *italic*.

Variable	Gain	Gain [dBi]
Antenna No.1	20.3476	13.0851
Antenna No.2	21.3475	13.2935
Antenna No.3	*21.6401*	*13.3526*
Antenna No.4	20.0586	13.023
Antenna No.5	19.1266	12.8161
Antenna No.6	17.6674	12.4717
Antenna No.7	15.1044	11.791

Tab. 3 Values of total gain.

Fig. 4 shows the dependence between total gain of simulated antennas and reflector diameter change. The maximum value of total gain was obtained for antenna No.3 with the value of 13.3526 dBi

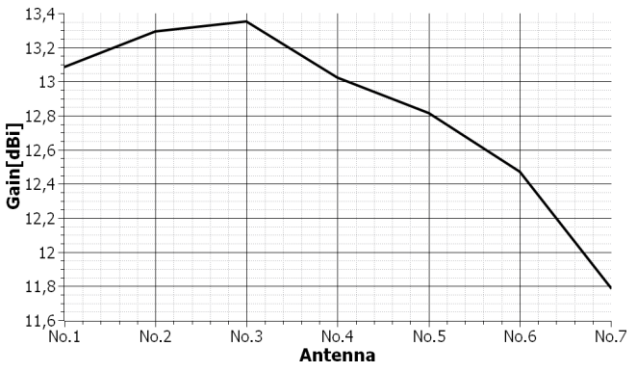


Fig. 4 Total gain of designed antenna operating on 5.2 GHz in dBi units.

The most important parameter of antenna is Half Power of Beam Width (HPBW). The HPBW of designed antennas is shown in Fig. 5.

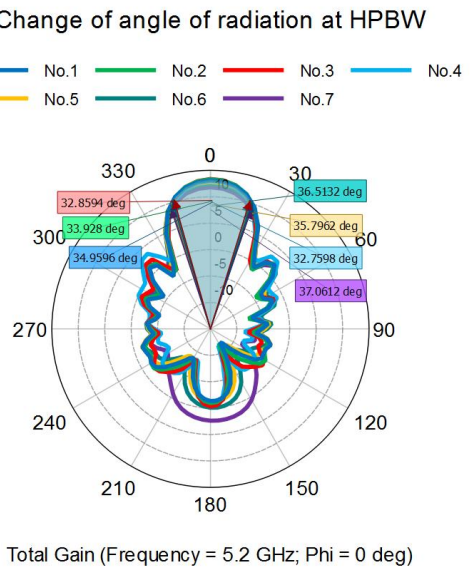


Fig. 5 Change of angle of radiation for $\Phi = 0°$.

Obtained results of change of radiation angle on HPBW is in Tab. 4. In this table, maximal values of obtained gain are denoted in *italic*. Antenna No.4 has the lowest value of radiation angle.

Variable	HPBW - $\Phi = 0°$	HPBW - $\Phi = 90°$
Antenna No.1	34.9596°	34.856°
Antenna No.2	33.928°	33.8198°
Antenna No.3	32.8594°	32.7449°
Antenna No.4	*32.7598°*	*32.5752°*
Antenna No.5	35.7962°	35.6773°
Antenna No.6	36.5132°	36.3535°
Antenna No.7	37.0612°	36.8921°

Tab. 4 Values of change of radiation angle.

Fig. 6 shows the dependency between the reflector diameter of helix antenna and obtained values of radiation angle. The curve illustrated by solid line represents values obtained on $\Phi = 0°$ and the dashed line represents values obtained on $\Phi = 90°$.

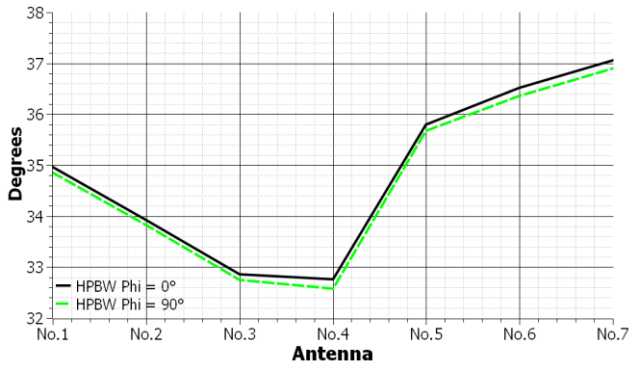


Fig. 6 Angle of radiation at HPBW (Half Power of Beam Width).

The obtained results of maximal radiated power is in the Tab. 5. In this table, the maximal values of obtained gain are denoted in *italic*. Antenna No.5 has the lowest value of power.

Variable	Power [mW]	Power [dBW]	SLL [dB]
Antenna No.1	4.95197	-23.0522	10.74
Antenna No.2	4.9928	-23.0166	*11.3573*
Antenna No.3	4.95596	-23.0487	11.1156
Antenna No.4	4.88396	-23.1123	9.47339
Antenna No.5	*5.03702*	*-22.9783*	10.5252
Antenna No.6	4.98735	-23.0213	10.1055
Antenna No.7	4.91923	-23.081	9.37469

Tab. 5 Values of obtained power and SLL (Side Lobe Level).

Fig. 7 shows the radiated power as function of reflector diameter (7 types of helix antennas).

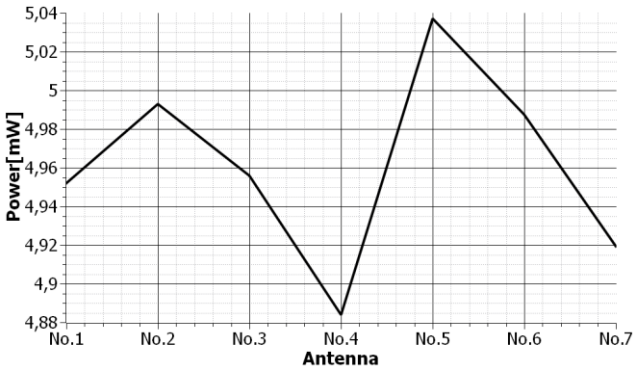


Fig. 7 Radiated power of designed antennas.

The important parameter of antenna is impedance. The placement of impedance points in Smith chart for each designed antenna is in Fig. 8.

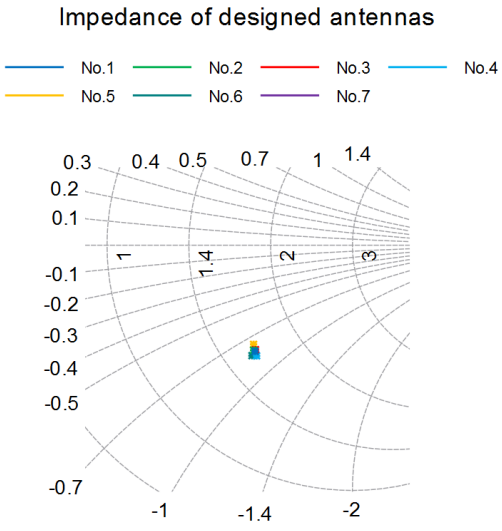


Fig. 8 The impedance of modeled antennas.

Conclusion

This paper is focused on examination of impact of reflector diameter of helix antenna. We are looking for appropriate antenna for FSO/RF hybrid systems. For our design, we chose the frequency of 5.2 GHz from non-licensed frequency range. Antenna No.3 has the largest value of gain: 21.6401 (13.3526 dBi). The HPBW of antenna No.4 for $\Phi = 0°$ is 32.7598° and 32.5752° for $\Phi = 90°$. For non-licensed application the maximal value of radiated power into affected area is very important parameter. The threshold value for the given frequency (5.2 GHz) is 20 mW. The maximal value of radiated power was obtained for antenna No.5 with the value of 5.03702 mW (-22.9783 dBW). The last examined parameter was side lobe level (SLL) which shows the ratio between the main lobe and the second maximal lobe. With regards on this knowledge the maximal value for antenna No.2 was 11.3573 dB. This paper shows that the trade-off is the need.

Acknowledgements

Research described in the paper was supervised by doc. Ľuboš Ovseník, KEMT, FEI, TUKE in Košice and supported by following research grants: VEGA 1/0772/17, KEGA 023TUKE-4/2017 and APVV-17-0208 - Resilient mobile networks for content delivery.

References

[1] MÁRTON, M., OVSENÍK, Ľ, ŠPES, M., "Measurement effect of visibility in experimental FSO system" In: *Informatics 2017*. - Danvers: IEEE, p. 249-252. - ISBN 978-1-5386-0888-3, 2017

[2] MÁRTON, M., OVSENÍK, Ľ, TURÁN, J., ŠPES, M., "Design and analysis of microstrip antenna for 2.46GHz in Program Suite FEKO" In: *Carpathian Journal of Electronic and Computer Engineering*. Vol. 10, no. 1, p. 31-34, ISSN 1844 – 9689, 2017

[3] TÓTH, J., OVSENÍK, Ľ., TURÁN, J., MICHAELI, L., MÁRTON, M., "Classification prediction analysis of RSSI parameter in hard switching process for FSO/RF systems", In: *Measurement*, vol. 116, pp. 602-610, 2018, 10.1016/j.measurement.2017.11.044

[4] MÁRTON, M., OVSENÍK, Ľ, HUSZANÍK, T., ŠPES, M., "Analysis of possibilities for measurement effect of visibility in experimental FSO system", Open Computer Science, vol. 8, no. 1, pp. 135-141, 2018, 10.1515/comp-2018-0013

[5] MÁRTON, M., OVSENÍK, Ľ, TURÁN, J., ŠPES, M., "Design of helix antenna for 2.4GHz applicable in FSORF hybrid system" In: MIPRO 2018, Rijeka, p. 469-472., ISBN 978-953-233-096-0 2018

[6] TÓTH, J., OVSENÍK, Ľ, TURÁN, J., "Free space optics experimental system - long term measurements and analysis", In: *Acta Electrotechnica et Informatica*. 2015. Vol. 15, no 2 (2015), pp.26-30

[7] DAVIDSON, D., B., THERON, I., P., JAKOBUS, U., LANDSTORFER, F., M., MEYER, F., J., C., MOSTER, J., TONDER, J., J., "*Communications and Signal Processing"*, 1998. 430 p. ISBN: 978-0-7803-5054-5

[8] CUTSHALL, R., T., ZIOLKOWSKI, R., W., "Performance Characteristics of Planar and Three-Dimensional Versions of a Frequency-Agile Electrically Small Antenna", In: *IEEE Antennas and Propagation Magazine*, vol. 56, pp. 53-71, 2014, ISSN 1045-9243

[9] SHEHAN, J., W., ADAMS, R., S., "Electrically thin magnetodielectric metamaterial fresnel lens antenna", *Proc. of IEEE International Symposium on Antennas and Propagation (APSURSI)* 2016, pp. 1387-1388, ISSN 1947-1491

[10] VOLAKIS, J., "*Antenna engineering handbook*", 1st ed. London: McGraw-Hill, 2007.

[11] MONDAL, J., KUMAR RAY, S., ALAM, M., RAHMAN, M., "Design Smart Antenna by Microstrip Patch Antenna Array", In: *International Journal of Engineering and Technology*, vol. 3, no. 6, pp. 675-683, 2011.

[12] BANSAL, L., PARKASH, D., "Design of Timer Shaped Microstrip-Fed Antenna for RFID Applications", In: *International Journal of Engineering Research*, vol. 3, no. 4, pp. 294-298, 2014.

[13] BLAKE, L., LONG, M.: *Antennas*, Raleigh, NC: SciTech Pub., 2009.

[14] KLINKENBUSCH, L., "Time domain near-field to near-field transformation using a spherical-multipole approach", In: *Radio Science*, vol. 46, no. 5, 2011.

[15] IEEE Standard 211-1997, „The IEEE Standard Definitions of Terms for Radio Wave Propagation", December 1997, 27 p.

[16] CARR, J., HIPPISLEY, G., "*Practical antenna handbook*." New York: McGraw-Hill, 2014.

About Authors...

Michal MÁRTON (Ing.) received Ing. (MSc.) degree in 2016 at Department of Electronics and Multimedia Telecommunications, Faculty of Electrical Engineering and Informatics of Technical University of Košice. Since September 2016 he has been at University of Technology, Košice as PhD. student. His research interests include optical fiber gyroscopic systems and optical communication systems.

Michal ŠPES (Ing.) was born in 1991. In 2015 graduated (MSc) at the department of electric power engineering. At present is a Ph.D. student in the Department of Electric Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in electric power engineering on subject evaluation of generator exciting outage. His scientific research is mainly focused on research of powerline ampacity system.

Jakub URBANSKÝ (Ing.) was born in 1991. In 2017 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. At present is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in electric power engineering on subject of the integration of electric vehicle into the electrical power distribution, using the vehicles-to-grid (V2G) technology. His scientific research is mainly focused on research of renewable energy sources.

Binarization of Feature Vectors for Image Retrieval

Erik OTTLIK[1], *Abin JOSE*[1]

[1]Institute of Communications Engineering, RWTH Aachen University, 52056 Aachen, Germany

erik.ottlik@rwth-aachen.de, jose@ient.rwth-aachen.de

Abstract. *Due to the vast size of image datasets nowadays, image retrieval has to be not only effective but also efficient. To cope with this, we can assign short binary codes to the images, which capture the similarity information. By using binary codes, one can quickly compute Hamming distances to retrieve the most similar data items. Recent deep hashing methods use a pairwise or triplet based loss for generating a feature space that captures the similarity information and is suitable for binarization. However, these loss functions cannot capture the semantic relationship between all training images since only pairs of images guide the feature space generation. A feature generation based on Deep Linear Discriminant Analysis (DeepLDA) [4] can better reflect the structure of the training data by the estimation of covariance matrices. First, we show that DeepLDA can be solved by the more general Deep Canonical Correlation Analysis (DCCA) [1] to generate a suitable feature space. Secondly, we combine our feature generation with Iterative Quantization (ITQ) [5] to produce the final binary codes. An analysis on CIFAR-10 shows that our method can outperform the state-of-the-art techniques for image retrieval.*

Keywords

Content-based image retrieval, Deep learning, Deep supervised hashing.

1. Introduction

In content-based image retrieval, we try to retrieve the images most similar to a query image. For describing the content of the images, we need to derive a feature space that captures the similarity information among the images.

The state-of-the-art deep hashing methods use Convolutional Neural Networks (CNNs) as feature extractors and combine them with binary coding procedures [7]. The resulting binary codes can be used to efficiently query for similar data items by an exhaustive Hamming distance computation or a hash table lookup.

Since the pairwise-based loss calculations cannot capture the similarity information among all training images simultaneously, it is arguable if the resulting feature space is optimal. For that reason, we discuss DeepLDA which has the desirable properties that only images belonging to the same class are close in the generated feature space [4].

1.1. Deep Linear Discrimant Analysis

In linear discriminant analysis (LDA), we try to find optimal linear projections $\mathbf{a}$ which maximize the between-class scatter $\mathbf{S}_B$ while minimizing the within-class scatter $\mathbf{S}_W$. The optimal linear projections must satisfy the eigenvalue problem given in (1). The best linear projections are the eigenvectors corresponding to the largest eigenvalues.

$$\mathbf{S}_W^{-1}\mathbf{S}_B\mathbf{a} = \lambda\mathbf{a}, \quad \text{Loss} = -\sum_i \lambda_i \tag{1}$$

Large eigenvalues indicate a good separation of the classes in the corresponding dimension. DeepLDA tries to push the eigenvalues as the loss of a CNN [4]. Since the eigenvalues are not bounded, DeepLDA tends to produce trivial solutions if the sum of all eigenvalues is maximized [4]. In particular, the network tends to push one eigenvalue to a very large value while the others stay close to zero. This results in a good separation in only one dimension. Dorfer et al. proposed to increase only the smallest eigenvalues to achieve a better separation in every dimension [4].

1.2. Deep Canonical Correlation Analysis

Hotelling's Canonical Correlation Analysis (CCA) [6] is closely related to LDA. CCA aims at finding optimal linear projections $\mathbf{a}$ and $\mathbf{b}$ which maximize the correlation ρ between two datasets $\mathbf{X}$ and $\mathbf{Y}$.

$$\rho = \max_{\mathbf{a},\mathbf{b}} \text{corr}(\mathbf{a}^T\mathbf{X}, \mathbf{b}^T\mathbf{Y}), \quad \text{Loss} = -\sum_i \rho_i \tag{2}$$

Andrew et al. proposed to correlate the output of two multilayer perceptrons through the CCA objective [1]. Their proposed Deep CCA (DCCA) loss is the negative sum of the correlation coefficients as given in (2) [1].

2. Solving DeepLDA by DCCA

Bartlett proved in 1938 that LDA can be solved by the more general CCA in a special setup [2]. Our idea is that we can also solve the DeepLDA objective by DCCA. We propose to correlate the output $\mathbf{X}$ of a CNN with a group indicator matrix $\mathbf{Y}$, which is constructed as given in (3).

$$\mathbf{Y}_{ij} = \begin{cases} 1, & \text{if image i belongs to class j} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

In the case of mutually exclusive classes, i.e., each image is labeled by a single concept, we can show that our setup solves the DeepLDA objective. The loss in our setup can be written as given in (4). For single-labeled images, we can express the loss in terms of the eigenvalues λ of the DeepLDA loss. We can directly infer that we maximize the sum of the eigenvalues by maximizing the correlation coefficients in our setup. The correlation coefficients are always bounded by one. Thus, we do not produce trivial solutions as in the LDA-based loss. Once we achieve a good separation in one dimension, the correlation coefficient becomes close to one. Then, the network has to increase the correlations coefficients in the other dimensions to decrease the loss further. This results in a good separation in every dimension.

$$\text{Loss} = -\sum_i \rho_i \overset{\text{for LDA equiv.}}{=} -\sum_i \sqrt{\frac{\lambda_i}{1+\lambda_i}} \tag{4}$$

3. Combination with ITQ

To generate the final binary codes we use Iterative Quantization (ITQ) [5]. It can be interpreted as fitting an m-dimensional hypercube to our m-dimensional feature vectors such that the quantization loss is minimized [5]. We apply ITQ on the feature vectors of all gallery images since ITQ is an unsupervised technique. The number of bits generated by ITQ corresponds to the dimensionality of our feature vectors. Their dimensionality is, in turn, bounded by the number of classes due to the properties of LDA and CCA. Therefore, we use an ensemble technique and concatenate the bits of different training runs with random initializations to increase the bit lenghts if necessary.

4. Results

We follow the common benchmark setup for CIFAR-10 [8]. We sample 1,000 query images. The remaining 59,000 images constitute the gallery dataset. From this gallery dataset, we sample 5,000 training images for finetuning our feature extraction.

In Tab. 1, we present the mean average precision (mAP) scores for our method and recent state-of-the-art methods.

Method	CIFAR-10			
	12 bits	24 bits	32 bits	48 bits
Ours	**0.7936**	**0.8306**	**0.8414**	**0.8512**
DDSH [7]	0.7695	0.8289	0.8352	0.8194
DSDH [9]	0.7442	0.7868	0.7991	0.8142
DTSH [11]	0.710	0.750	0.765	0.774
DPSH [10]	0.6844	0.7225	0.7396	0.7460

Tab. 1. mAP on CIFAR-10 dataset. The scores of DDSH, DSDH, and DPSH are cited from [7]. The scores of DTSH are cited from [11]. The same experimental setup was used for all methods (see [11]). All methods use CNN-F [3] as feature extractor.

Our framework can project to a maximally 9-dimensional space in the case of CIFAR-10 due to the LDA and CCA boundaries. The mAP score of our method for 9 bits, so without the ensemble technique, is 0.7796.

5. Conclusions

Firstly, we showed that DeepLDA can be regarded as a special case of DCCA. Secondly, our proposed framework exploits this equivalence to generate an effective feature space which can be easily binarized by ITQ. The conducted benchmarks on CIFAR-10 indicate that our method can outperform the state-of-the-art techniques for content-based image retrieval.

References

[1] ANDREW, G., ARORA, R., BILMES, J., LIVESCU, K. Deep canonical correlation analysis. In *International Conference on Machine Learning*, pages 1247–1255, 2013.

[2] BARTLETT, M. Further aspects of the theory of multiple regression. *Mathematical Proceedings of the Cambridge Philosophical Society*, 34(1):33–40, 1938.

[3] CHATFIELD, K., SIMONYAN, K., VEDALDI, A., ZISSERMAN, A. Return of the devil in the details: Delving deep into convolutional nets. *arXiv preprint arXiv:1405.3531*, 2014.

[4] DORFER, M., KELZ, R., WIDMER, G. Deep linear discriminant analysis. *arXiv preprint arXiv:1511.04707*, 2015.

[5] GONG, Y. LAZEBNIK, S., GORDO, A., PERRONNIN, F. Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12):2916–2929, 2013.

[6] HOTELLING, H. Relations between two sets of variates. *Biometrika*, 28(3/4):321–377, 1936.

[7] JIANG, Q., CUI, X., LI, W. Deep discrete supervised hashing. *IEEE Transactions on Image Processing*, 27(12):5996–6009, 2018.

[8] KRIZHEVSKY A., HINTON G. Learning multiple layers of features from tiny images. Technical report, Citeseer, 2009.

[9] LI, Q., SUN, Z., HE, R., TAN, T. Deep supervised discrete hashing. In *Advances in Neural Information Processing Systems*, pages 2482–2491, 2017.

[10] LI, W., WANG, S., KANG, W. Feature learning based deep supervised hashing with pairwise labels. *arXiv preprint arXiv:1511.03855*, 2015.

[11] WANG, X. SHI, Y., KITANI, K. Deep supervised hashing with triplet labels. In *Asian Conference on Computer Vision*, pages 70–84. Springer, 2016.

Hierarchical Depth Map Estimation for Inter-Prediction in Video Coding

Ishan SAXENA, Hossein BAKHSHI GOLESTANI

Institute of Communications Engineering, RWTH Aachen University, 52056 Aachen, Germany

ishan.saxena@rwth-aachen.de, golestani@ient.rwth-aachen.de

Abstract. *In this paper, we introduce two hierarchical structure based methods for estimating depth maps that will iteratively reduce the inter-camera baseline between reference pictures. The methods further utilize the 3D scene geometry and camera motion information to improve the estimation of depth maps leading to the synthesis of better reference pictures using depth image-based rendering techniques. The final synthesized reference pictures for the video sequence are ultimately provided as an additional reference picture to the High Efficiency Video Coding (HEVC) encoder, that aids in improving its motion compensation performance. Using our proposed methods, we achieve a bit-rate reduction of 7.38% on average, compared to HEVC.*

Keywords

3D scene geometry, camera motion, depth map, HEVC, hierarchical structure, motion compensation.

1. Introduction

Video content consumption is increasing by the day and is projected to reach 82% of all internet traffic by 2022 [1]. Add to this growth spurt the increased affinity of users towards higher resolution content on multiple devices which has made it even more crucial to compress the content more efficiently without any significant loss of quality. Hence, there is a strong need to incorporate further improvements to current video compression standards such as HEVC [2].

Inter-prediction in video compression relies on utilizing the temporal redundancy present in previously decoded pictures to predict the current picture more efficiently using Motion Compensation (MC). The previously decoded pictures act as a reference for prediction of subsequent pictures and are called Reference Pictures. In HEVC, usually two sets of Reference Picture Lists (RPLs) are utilized for MC. The proposed work relies upon generating an Additional Reference (AR) for MC using information about 3D structure of the scene and prior knowledge of camera motion. This technique for generating ARs has been previously presented in [3]-[5]. In [3], the 3D model of a scene is reconstructed by using a sub-set of the Key-Frames (KFs) and updated continuously as more KFs become available at the decoder. The ARs are then synthesized by re-projection of the 3D model onto a 2D plane using the camera parameters which are obtained from applying Structure-from-Motion (SfM) algorithm on all available pictures at the decoder. In [4], the authors estimate the ARs using an inverse 3D warping view synthesis process that requires depth maps and texture information from the two nearest KFs. The depth maps are estimated using a multi-view stereo process using all KFs present at the HEVC Picture buffer. Most recently, in [5], the authors extended the previous work and proposed using B-Frames (BFs) in addition to KFs as reference views for synthesizing ARs. The depth maps for each reference picture were again generated, similar to the previous work, using all available KFs at the HEVC Picture buffer.

This work follows the basic process flow used in these previous works, while using a two-view hierarchical depth map estimation approach instead of a multi-view one. The poor performance of multi-view approaches in texture-less regions of pictures and the absence of local texture details in estimated depth maps convinced us to use a two-view approach in this work. It was further observed that fast motion of recording camera through a scene results in a high inter-camera baseline between consecutive KFs of the sequence. Having such a high baseline between reference pictures is detrimental to the quality of the estimated depth map. Hence, we propose to use reference pictures that have a lower baseline in the current work.

In the first method, these low-baseline references are determined using a hierarchical structure derived from the HEVC decoding order. The second method additionally updates the depth maps estimated for current reference pictures using subsequently decoded pictures, This leads to synthesis of improved ARs for future pictures. The ARs synthesized from depth maps estimated using the proposed methods help improve the coding efficiency of the HEVC encoder resulting in a bit-rate reduction of 7.27% for the first method and 7.48% for the second method, on average, over HEVC and further surpassing results obtained by the previously mentioned works.

The rest of this paper is structured as follows. Section 2 describes the two proposed methods for depth map esti-

mation. Section 3 discusses the simulation results obtained from using the proposed methods on various test sequences while Section 4 provides the conclusions.

2. Proposed Methods

The process flow utilized in the current work is presented in Fig. 1. The 3D geometry of the scene is derived by applying Structure-from-Motion (SfM) [3] on all the input video pictures present in a sequence of length M. The camera parameters $\mathbf{CP} = \{CP_i \mid i = 1, 2, ..., M\}$ and the point cloud of the scene are provided at the output of the SfM algorithm. The extremities of the point cloud, as seen from each camera position, determine the disparity range $\mathbf{Z} = \{Z_i \mid i = 1, 2, ..., M\}$ over which the depth map is coded. For estimating depth maps, we use Depth Estimation Reference Software (DERS) [6] in this paper. It is based on the Graph-cut depth estimation approach and provides flexibility for choosing the number of input views, matching methods, depth estimation modes, etc. depending upon the intended application. The specific images $\mathbf{F} = \{F_i \mid i = 1, 2, ..., M\}$ provided as reference inputs are determined using the proposed methods.

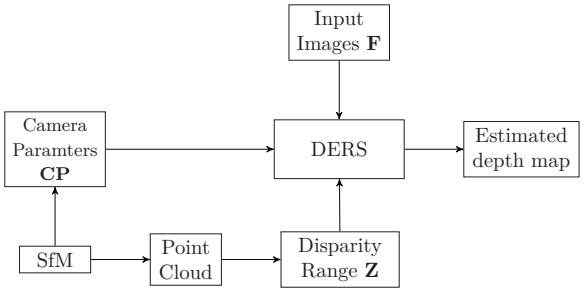


Fig. 1: Block diagram showing the proposed process flow for depth map estimation.

2.1. Conventional Method (CM)

The Conventional Method (CM) uses the two nearest KFs as reference inputs for estimating depth maps of all intermediate pictures present inside a Group of Pictures (GOP). However, these references have a high inter-camera baseline (equal to GOP size or G) and hence the depth maps estimated suffer from occlusions and matching errors leading to a degradation in their quality.

2.2. Hierarchical Structure (HS)

The proposed methods try to exploit the fact that an iterative reduction in inter-camera baseline between reference pictures leads to a reduction in occlusion and matching errors in the resultant estimated depth map. This reduction is proposed to be achieved by using reference pictures according to the HS as shown in Fig. 2. The lower Temporal ID (TID) level pictures already available in the Picture buffer are used as reference pictures while generating depth maps for future (or higher TID level) pictures. However, lowering the inter-camera baseline also impacts accuracy of the estimated depth map. Hence, the camera baseline should only be reduced to such an extent where the overall depth map quality doesn't degrade. Algorithms 1 and 2 provide an algorithmic explanation for choosing reference pictures for KFs and BFs respectively. Additional notations include P which refers to Intra Random Access Point (IRAP). DM_i refers to the estimated depthmap for the i^{th} picture. KF is a set consisting of index values for KFs in $\mathbf{F}$ and BF is a set consisting of index values for all BFs in $\mathbf{F}$.

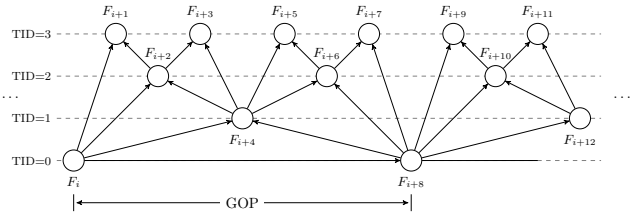


Fig. 2: Proposed HS with arrows emanating from reference pictures towards future pictures.

Algorithm 1 Estimation of depth maps for KFs in HS

Require: $\mathbf{F}, \mathbf{CP}, \mathbf{Z}, M, G, P$
For all pictures F_i where $i \in KF$
if $mod(i-1, P) == 0$ **then**
estimate DM_i using F_{i+G}, CP_i and Z_i
else
estimate DM_i using F_{i-G}, CP_i and Z_i
end if
return DM_i

Algorithm 2 Estimation of depth maps for BFs in HS

Require: $\mathbf{F}, \mathbf{CP}, \mathbf{Z}, M, G, P$
For all pictures F_j where $j \in BF$
if $mod(j, 2) == 0$ **then**
estimate DM_j using F_{j+1}, F_{j-1}, CP_j and Z_j
else if $mod(j, G/2) == (G/2 - 1)$ **then**
estimate DM_j using $F_{j+G/4}$, $F_{j-G/4}$, CP_j and Z_j
else if $mod(j, G/2) == (G/4 - 1)$ **then**
estimate DM_j using $F_{j+G/2}$, $F_{j-G/2}$, CP_j and Z_j
end if
return DM_j

A short explanation for how the process happens for a sample GOP (assuming $G = 8$ as shown in Fig. 2) is provided below.

- At first, only the KFs are transmitted serially to the HEVC encoder. For these pictures, only the previous KF is available to act as a reference for depth map estimation. A special case exists for KFs occurring just

after P, where the previous KF can not be used as a reference, hence we propose to use the next KF as reference. When the first picture F_i of the sample GOP reaches the HEVC encoder, F_{i-8} acts as a reference to estimate DM_i. Similarly for F_{i+8}, F_i acts as a reference to estimate DM_{i+8}. For the case of KFs, since only one picture is used as a reference and the inter-camera baseline is high (= G), the quality of depth map estimated is poor.

- Next, to estimate depth map for BFs in TID = 1 (F_{i+4}), we have both KFs available in the GOP to act as reference pictures. The inter-camera baseline is still high (= G), but since we have two separate reference views, the occlusion errors will be less which leads to an improvement in quality of the estimated depth map (DM_{i+4}). It should be noted that till this step, the algorithm is working in exactly the same way as the conventional methods which use only KFs for depth map estimation.

- Traversing higher into the HS, we next estimate depth map for BFs in TID = 2. Taking the case of F_{i+2} as an example, we have 3 pictures in the GOP that can act as references, F_i, F_{i+4} and F_{i+8}. The inter-camera baseline is minimized by using the closest pictures available, which are F_i and F_{i+4} and hence these two are used to estimate DM_{i+2}.

- Finally arriving at the last TID = 3 level for the sample GOP, we again try to estimate depth maps using the lowest possible baseline. For eg., in the case of F_{i+1}, the two closest reference pictures will be F_i and F_{i+2}, which are then used to estimate DM_{i+1}.

The same process can be repeated for each GOP of the sequence.

2.3. Hierarchical Structure plus updated depth maps (HS+UD)

As observed from Fig. 2, KFs and TID = 1 BFs act as reference pictures multiple times for synthesizing future pictures. However, their depth maps are not optimal in quality since they are estimated using previous KFs with a high inter-camera baseline. Hence, BFs from higher TID levels, once they are available in the Picture buffer, are used as references and the depth maps of KFs and TID = 1 BFs are updated. This improves the synthesis results of future pictures that will use depth maps of these lower TID pictures as references. Updated depth maps are annotated as DM_i', DM_i'', DM_i''' and so on, depending upon the number of times they have been updated. This method is represented algorithmically in Algorithms 3, 4 and 5. HS+UD still relies on the HS method for estimation of initial KF depth maps.

Algorithm 3 Estimation of updated depth maps after availability of TID = 1 pictures in buffer

Require: $\mathbf{F}, \mathbf{CP}, \mathbf{Z}, M, G, P$
  {Update $DM_i \rightarrow DM_i'$ for pictures in TID = 0}
  if $mod(i, G) == 1$ **and** $mod(i, P) == 1$ **then**
    estimate DM_i' using $F_{i+G/2}$, CP_i and Z_i
  else if $mod(i, G) == 1$ **then**
    estimate DM_i' using $F_{i-G/2}$, CP_i and Z_i
  end if
  return DM_i'

Algorithm 4 Estimation of updated depth maps after availability of TID = 2 pictures in buffer

Require: $\mathbf{F}, \mathbf{CP}, \mathbf{Z}, M, G, P$
  {Update $DM_i' \rightarrow DM_i''$ for pictures in TID = 0 and $DM_i \rightarrow DM_i'$ in TID = 1}
  if $mod(i, G) == 1$ **and** $mod(i, P) == 1$ **then**
    estimate DM_i'' using $F_{i+G/4}$, CP_i and Z_i
  else if $mod(i, G) == 1$ **then**
    estimate DM_i'' using $F_{i-G/4}$, CP_i and Z_i
  else if $mod(j, G/2) == (G/4 - 1)$ **then**
    estimate DM_i' using $F_{i+G/2}$, $F_{i-G/2}$, CP_i and Z_i
  end if
  return DM_i', DM_i''

Algorithm 5 Estimation of updated depth maps after availability of TID = 3 pictures in buffer

Require: $\mathbf{F}, \mathbf{CP}, \mathbf{Z}, M, G, P$
  {Update $DM_i'' \rightarrow DM_i'''$ for pictures in TID = 0}
  if $mod(i, G) == 1$ **and** $mod(i, P) == 1$ **then**
    estimate DM_i''' using $F_{i+G/4}$, $F_{i+G/4-1}$, CP_i and Z_i
  else if $mod(i, G) == 1$ **then**
    estimate DM_i''' using $F_{i-G/4}$, $F_{i-G/4+1}$, CP_i and Z_i
  end if
  return DM_i'''

An example is presented below for the first GOP of a sample sequence (G = 8) to better understand the exact steps followed in this method.

- Initially the depth maps for KFs (DM_0 and DM_8) are estimated using the previously available KFs in the Picture buffer, similar to the HS method. Next, these two KFs serve as references for estimating the depth map DM_4 for the picture in TID = 1.

- Now, since the picture F_4 is also available in the Picture buffer, in addition to being used for estimating the depth maps of future pictures, it can also be used for updating the depth maps of F_0 and F_8.

- In the next step, depth maps for pictures in TID = 2 are estimated.

- At this stage, both the pictures in TID = 2 are also available to serve as references for updating the depth maps for pictures in TID = 0 and 1.
- The next step consists of estimating depth maps for pictures in TID = 3 using the closest reference pictures available according to the HS.
- Finally, the pictures in TID = 3 can serve as references for the final depth map updates of F_0 and F_8. A similar procedure is followed for the next GOPs too.

3. Simulation Results and Discussion

The actual depth maps were estimated based on HS and HS+UD using the DERS version 6.1 [6]. The estimated depth maps were utilized to synthesize virtual views using the View Synthesis Reference Software (VSRS) [7]. The synthesized views were added as ARs at the last position for both RPLs (List0 and List1) and encoding results were simulated using HEVC Test Model (HM) 16.7. Since we are just adding ARs to the RPLs and not replacing the already available ones generated by the HEVC encoder, we can guarantee at least base performance of the encoder.

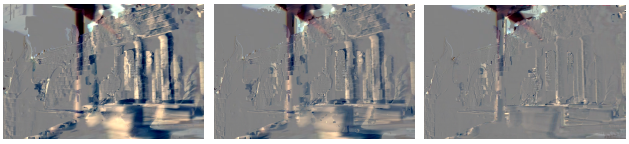
The proposed methods were tested on 5 Full-HD sequences including Sintel, GTFly, IndianBuilding1, DayLightRoad and IceRock2. A QP range of 25, 29, 33 and 37 were used for all sequences except IndianBuilding1 for which a QP range of 21, 25, 29 and 33 was used. The GOP size used was 8 while the IRAP was 32.

A visual comparison of the performance of CM, HS and HS+UD is provided in Fig. 3. This figure compares the prediction errors (defined as the absolute difference between the ground-truth and predicted picture) for CM, HS and HS+UD. The lower prediction errors obtained for HS and HS+UD highlight that reduction in inter-camera baseline leads to an improvement in quality of depth maps further resulting in better predicted AR. Furthermore, the quantitative results for improvement in BD-Rate [8], BD-YPSNR and actual usage rate of AR (defined as the fraction of total MC prediction done by the HEVC encoder using provided AR) were compared against CM and the 3D reconstruction methods: 3D [4] and 3DH [5]. As observed from Tables 2 and 1, HS and HS+UD dominate over all other methods for both metrics on average, thus validating the proposed hypothesis. The use of depth maps estimated using HS for synthesizing ARs leads to a 7.27% bit-rate reduction while the use of depth maps estimated from HS+UD leads to a 7.48% bit rate-reduction as compared to HM16.7.

For the test sequences, results obtained using HS+UD dominate over HS for each sequence except for the IceRock2 sequence, where absence of occluding objects in the scene and a low inter-camera baseline between KFs from the beginning leads to degradation of overall depth map quality when the baseline is further reduced. Additionally, as mentioned earlier, two-view methods tend to dominate over 3D reconstruction methods due to the presence of holes and absence of local textures in depth maps of 3D methods. However, for scenes such as Sintel, where the camera motion is in an almost semi-circular arc around an old building, where the redundant information derived from past and future pictures, as present in 3D reconstruction methods, overpowers these drawbacks leading to the estimation of improved depth maps and better bit-rate savings as shown in Tab. 2.

Tab. 1: Average contribution of AR in HEVC MC.

	Methods	AR Usage Rate (%)
Two-view	**CM**	12.16
	HS	24.85
	HS+UD	26.74
3D reconstruction	**3D**	12.64
	3DH	25.04



(a) (b) (c)

Fig. 3: Prediction error obtained from comparing ground truth with view synthesized using (a) CM, (b) HS and (c) HS+UD.

4. Conclusion

In this paper, we presented two methods for estimating depth maps for video sequence pictures using a hierarchical structure based on the HEVC encoding order. HS uses pictures already decoded and available in the HEVC Picture buffer as references to estimate the depth maps for current picture while HS+UD additionally updates the depth maps of previously available pictures using more recently available ones.

The improvements in BD-Rate, BD-YPSNR and increased usage rate of the generated AR by HEVC encoder for MC over other conventional and 3D based methods validates the good performance of the proposed methods. Future works can focus on fusion of depth maps estimated using hierarchical methods and 3D methods and optimizing the selection of reference pictures in HS+UD for cases where a lower inter-camera baseline degrades the depth map quality.

Tab. 2: Comparison of BD-Rate (%) and BD-YPSNR (dB) values for different sequences - Anchor: HEVC (HM16.7).

Sequences	Two-view						3D reconstruction			
	CM[1]		HS[2]		HS+UD[3]		3D[4]		3DH[5]	
	BD-Rate	BD-YPSNR	BD-Rate	BD-YPSNR	BD-Rate	BD-YPSNR	BD-Rate	BD-YPSNR	BD-Rate	BD-YPSNR
Sintel	-2.06	0.09	-3.20	0.14	-3.93	0.18	-4.23	0.19	**-5.36**	**0.24**
GTFly	-8.51	0.25	-11.90	0.35	**-12.24**	**0.36**	-7.74	0.23	-10.24	0.30
IndianBuilding1	-8.23	0.32	-10.90	0.42	**-10.97**	**0.43**	-8.49	0.33	-10.42	0.41
DayLightRoad	-2.28	0.08	-2.93	0.10	**-3.26**	**0.12**	-2.24	0.08	-2.38	0.08
IceRock2	-7.20	0.26	**-7.42**	**0.27**	-6.98	0.26	-6.09	0.22	-6.08	0.22
Average	-5.66	0.2	-7.27	0.26	**-7.48**	**0.27**	-5.76	0.21	-6.90	0.25

[1] Conventional Method
[2] Hierarchical Structure
[3] Hierarchical Structure plus updated depth maps
[4] 3D Reconstruction [4]
[5] 3D Reconstruction + Hierarchical [5]

References

[1] CISCO. Cisco Visual Networking Index Forecast and Trends 2017-2022. Nov. 2018.

[2] SULLIVAN, G. J., OHM, J., HAN, W., WIEGAND, T. Overview of the High Efficiency Video Coding (HEVC) Standard. In *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, pp. 1649-1668.

[3] GOLESTANI, H. B., SCHNEIDER, J., WIEN, M., OHM, J. Point cloud estimation for 3D structure-based frame prediction in video coding. In *Proceedings of the 2017 IEEE International Conference on Multimedia and Expo (ICME)*, Hong Kong, 2017, pp. 1267-1272.

[4] GOLESTANI, H. B., WIEN, M., OHM, J. 3D scene model based frame prediction in video coding. In *Proceedings of the 2017 International Conference on 3D Immersion (IC3D)*, Brussels, 2017, pp. 1-6.

[5] GOLESTANI, H. B., OHM, J. View Synthesis Prediction for Video Compression. In *Submitted to IEEE International Conference on Multimedia and Expo (ICME) 2019*.

[6] WEGNER, K., SENOH, T., LAFRUIT, G. Description of exploration experiments on freeviewpoint television (FTV). In *ISO/IEC JTC1/SC29/WG11 N 14551*, 2014.

[7] SENOH, T., Yamamoto, K., Tetsutani, N., Yasuda, H., Wegner, K. View Synthesis Reference Software (VSRS) 4.2 with improved inpainting and hole filing, In *ISO/IEC JTC1/SC29/WG11 MPEG2017, M40657*, Hobart, Australia, April 2017.

[8] BJONTEGAARD, G. Calculation of average PSNR differences between RD-curves. In *VCEG-M33*, 2001, USA.

About Authors...

Ishan SAXENA was born in Bhopal, India in 1994. He is currently pursuing his Masters Degree in Electrical Engineering, Information Technology and Computer Engineering from RWTH Aachen University.

Hossein BAKHSHI GOLESTANI was born in Mashhad, Iran. He received his Masters degree in Electrical Engineering from Sharif University of Technology, Tehran, Iran and is currently pursuing his Ph.D. Degree at the Institute of Communications Engineering, RWTH Aachen University, focusing on 3D scene models for frame prediction.

Modeling of a Hump Effect Using a Three-Dimensional TCAD Device Simulator

Dalibor BARRI[1,2]

[1]Dept. of Microelectronics, Czech Technical University, Technická 2, 166 27 Prague 6, Czech Republic
[2]STMicroelectronics, Pobřežní 620/3, 186 00, Prague 8, Czech Republic

barrridal@fel.cvut.cz

Abstract. *In this paper, we have analyzed the hump effect of polysilicon planar NMOS transistors. The hump effect caused by boron segregation makes a negative effect on transfer characteristic in the sub-threshold region of high voltage transistors. Therefore, we have analyzed the hump effect for several geometries, where various physical parameters have been lengths of MOS transistors, and gate thicknesses oxide of MOS transistors. Moreover, it has been studied for different simulation parameter. As the simulation parameters have been various gate-source voltages V_{GS}, and bulk-source voltage V_{BS}. In order to respect the hump effect in circuit level simulations, there has been verified a macromodel of the MOS transistor with the hump effect. All simulations have been run by the 3D TCAD Silvaco simulation tool.*

Keywords

MOS transistor, threshold voltage, shallow trench isolation, hump effect.

1. Introduction

In the present, there is a trend to use new design integrated circuits (ICs) in low power mode in the sub-threshold regime of MOS (Metal-Oxide-Semiconductor) transistors. The sub-threshold area has become as standard for low power applications due to very low current level with low supply voltage [1]. Moreover, the mainstream of the semiconductor industry is increasing the density of ICs on a chip. It is possible due to the evolution of submicron technologies, where the transition between active area and isolation becomes more abrupt to increase device integration density and performance. This is possible using shallow trench isolation (STI) process instead of a local oxidation of silicon (LOCOS) process. The STI process allows downscale lateral CMOS technology family and thus to reach a higher device density on a chip [2]. In case of a combination low-power and high-power devices in STI technology process option, the hump effect has been observed.

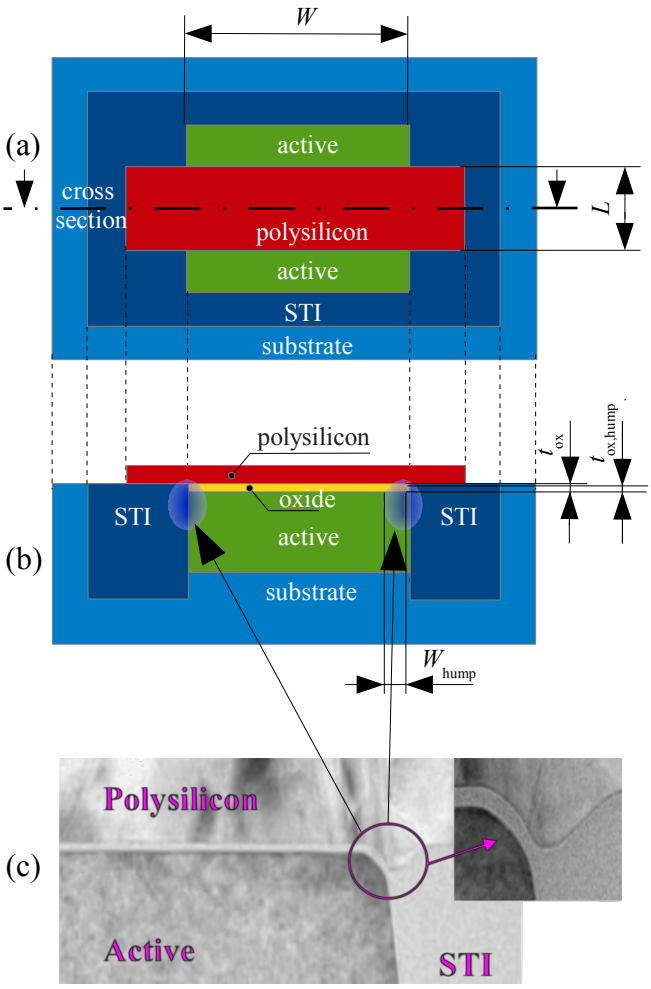


Fig. 1. Top view of the MOS transistor (a), its cross-section view (b) with highlighted parts of MOS transistor where the hump effect is located. The (c) part is TEM micrograph photo [3] of the fabricated MOS transistor.

Although the STI process offers to us a lot of improvements the abrupt separation of an active area of MOS transistors from STI oxide field (Fig. 1) fundamentally alters electrical characteristics of the MOS transistors. As one of the most important parameter of the MOS transistor, that is modified, is drain current I_D of MOS transistors in sub-threshold regime (see Fig. 3).

In this context, the main purpose of this article is to define constraints where the hump effect appears, because it isn't included in standard model cards of MOS devices with BSIM models. In order to be able to model the hump effect, the 3D TCAD simulation tool from Silvaco company [4] has been used. It is very important to understand the criteria, where the hump effect is modifying the drain current of MOS transistor, and to respect it in a low-power design.

2. A Hump Effect Description

The hump effect is a manufacturing defect of IC with STI or LOCOS option. This defect is visible in a transmission electron microscopy (TEM) cross-section in the width direction of the MOS transistor as it is shown in Fig. 1c. In these affected parts of the MOS transistors, the gate oxide thickness at the channel edge is lower then the gate oxide thickness at the channel center of the MOS transistors [3]. Moreover, there is also another aspect that has a negative impact on the channel. It is not a consistent concentration profile. Interface doping concentration between STI and substrate (at the channel edge of the MOS transistors) is decreased [5].

A decreasing both thickness gate oxide and concentration substrate profile at the edges of the MOS transistor results in a different threshold voltage V_{th} of these parts than it is in the rest of the MOS transistor. For these entire mechanisms and thinking symmetry properties of two parasitic corner devices, the hump effect is modeled as a composition of two different transistors. The first one is the main MOS transistor M_{main} and the other one is the hump transistor M_{hump}. The hump MOS transistor M_{hump} is located twice (on both sides of the width of the MOS transistor). Then, we can create a macro-model of the MOS transistor Fig. 2 with a hump effect, as suggested in [6].

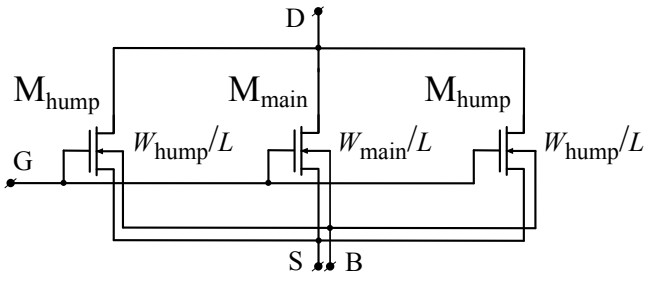


Fig. 2. Macro-model of hump effect on MOS transistors

2.1. Threshold Voltage V_{th}

Threshold voltage V_{th} plays a key role in semiconductor instances (such as LV MOS, HV MOS transistors, etc.), and any unpredictable behavior causes problems during front-end analog design. As has been mentioned in the previous section, parasitic hump devices present in standard MOS transistor cause a deviation of drain current I_D due to lower threshold voltage than it is in the rest of the MOS transistor. As a result of it and a partial goal of this paper is to be able to model the hump effect such as in the following Fig. 3

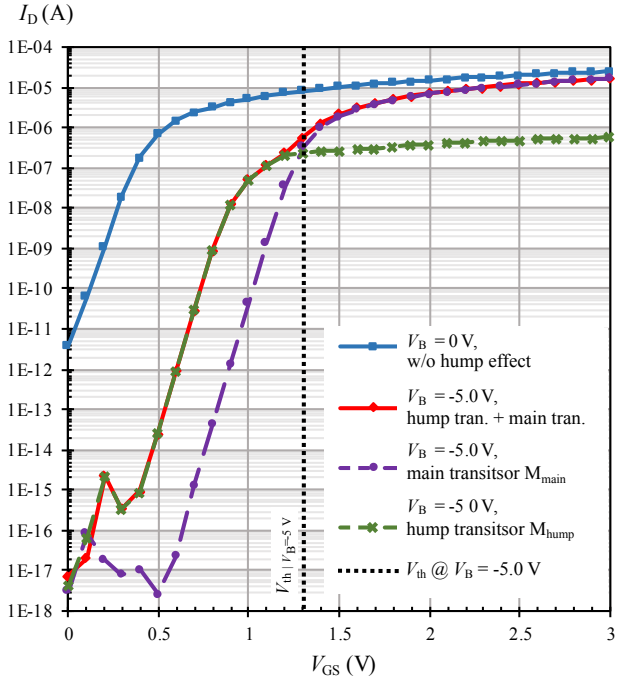


Fig. 3. Hump effect on I_D-V_{GS} characteristics of MOS transistors with diamond layout shape and with body effect V_B = -5.0 V

The threshold voltage V_{th} equals the sum of the flatband voltage V_{FB}, twice the bulk potential Φ_p and the voltage across the oxide due to the depletion layer. It is generally defined in [7] as the following expression:

$$V_{th} = V_{FB} + \gamma\sqrt{2(-\Phi_p)} - 2\Phi_p \tag{1}$$

where the V_{FB} is the flatband voltage, a Φ_p is workfunction of a substrate, and a γ is a body. After routine operation we will get the following expression with basic technology parameters:

$$V_{th} = \frac{T\log\left(\frac{N_A}{n_i}\right)k_B}{q} - \frac{T\log\left(\frac{N_D}{n_i}\right)k_B}{q} + \frac{2\sqrt{\frac{T\log\left(\frac{N_A}{n_i}\right)k_B}{q}}t_{ox}\sqrt{qN_A\varepsilon_{Si}}}{\varepsilon_{ox}} \tag{2}$$

where T is the temperature, k_B is the Boltzmann constant ($k_B = 1.38\cdot10^{-23}$ J/K), and q is the magnitude of the electron charge ($q = 1.602\cdot10^{19}$ C). N_A is the net ionized acceptor density, N_D is the donor impurity density in silicon, and n_i is intrinsic doping concentration of silicon ($n_i = 1.45\cdot10^{10}$ cm^{-3}). ε_{Si} is the permittivity of silicon, and ε_{ox} is the permittivity of oxide.

3. 3D TCAD Experiments

To be able to model hump effect, 3D TCAD simulations have been run for different settings of the MOS transistor. The list of basic parameters that have been various are described in Fig. 1 and listed below:

- L is a length of the MOS transistor;
- t_{ox} is a gate thickness oxide;
- V_B is a bulk voltage of the MOS transistor;
- V_{GS} is a gate-source voltage of the MOS transistor.

In this paper, we have analyzed an NMOS transistor that is shown in 3D visualization in Fig. 4 and Fig. 5. In this case, a width of the MOS transistor is $W = 5\,\mu m$, a width of the hump transistor is $W_{hump} = 0.5\,\mu m$, and a length of both MOS transistors is $L = 1\,\mu m$. The bulk-source voltage V_{BS} has been varied in a range from $V_{BS} = 0\,V$ to $V_{BS} = -5\,V$. The drain-source voltage V_{DS} has been set to $V_{DS} = 0.1\,V$. The oxide thicknesses have been set to $t_{ox} = 8.5\,nm$, and $t_{ox,hump} = 7.65\,nm$. The net ionized acceptor densities have been set to $N_A = 1.65 \cdot 10^{17}\,cm^{-3}$ for main MOS transistor and $N_{A,hump} = 1.65 \cdot 10^{16}\,cm^{-3}$ for the hump MOS transistor. The donor impurity density has been set to $N_D = 1.0 \cdot 10^{20}\,cm^{-3}$.

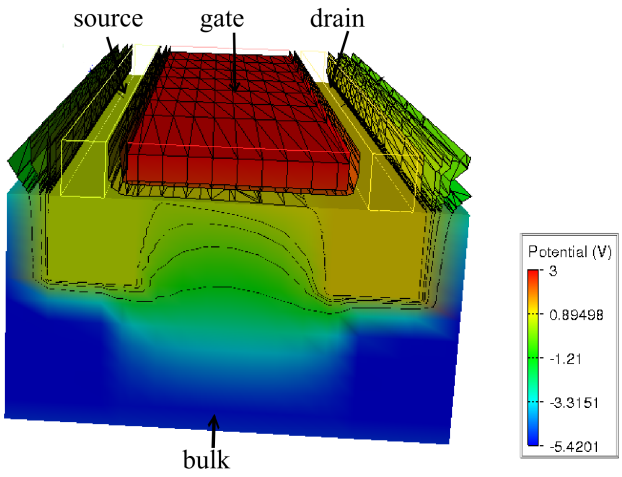


Fig. 4. An example of simulated MOS structure in 3D TCAD device simulator

In Fig. 5, there is a potential line across the MOS transistor for $V_{BS} = -0.2\,V$. As can be seen, the hump part (the front of the figure) has a lower voltage potential barrier that is not a barrier to the current compared to the main MOS transistor (the back of the figure). It affects the drain current I_D of the MOS transistor as shown in Fig. 3

In Fig. 6 and Fig. 7 we have explored the hump effect for three different cases. As the figure of merit parameter have been chosen current differences δI_{DS} (3) between a real MOS transistor with the hump effect $I_{DS,real}$ and an ideal MOS transistor $I_{DS,ideal}$.

$$\delta I_{DS} = I_{DS,real} - I_{DS,ideal} \tag{3}$$

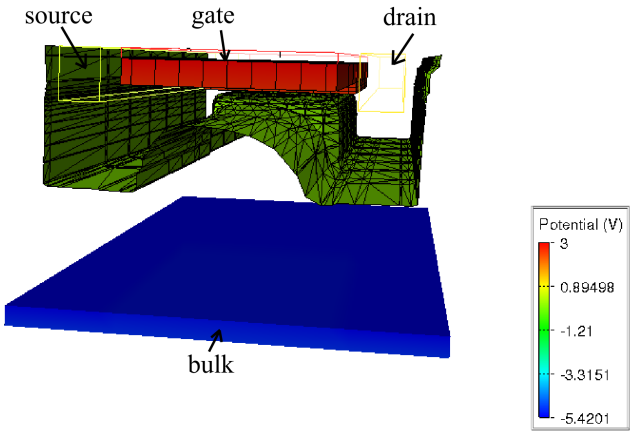


Fig. 5. Cross-section of the simulated MOS transistor with the hump effect

The first explored case (Fig. 6) of the hump dependence has been on the length of MOS transistor L. As we can see, the hump effect is present in the sub-threshold part of the MOS transistor and with the enlarging channel length of the MOS transistor, the hump effect also appears for higher gate-source voltage V_{GS}.

The second explored case of the hump effect has been the dependence of the MOS transistor on the gate oxide thickness. As we can see, the hump effect is more present for a thicker gate oxide thickness of the MOS transistor due to a shift of threshold voltage. The dependency of gate oxide thickness on final threshold voltage is present in 1.

4. Summary and conclusion

In this paper, we have observed the hump effect as a combination of two MOS transistors that are connected in a parallel connection. The first one is the main MOS transistor M_{main}, and the second part is a composition of two parasitic channel edge hump MOS transistors M_{hump} in parallel combination. The main and parasitic hump MOS transistor have different threshold voltage V_{th} and the parallel circuit reproduces very well the experimental transfer characteristic. The effect of the hump effect has been studied also for different lengths L of MOS transistor, for different bulk voltages V_B of MOS transistor, and for different gate oxide thicknesses t_{ox} of the MOS transistor.

In the next part of this paper, an analytical expression of the threshold voltage V_{th} of a general planar MOS transistor has been described. The analytical expression offers a very good overview of all parameters that influence threshold voltage V_{th}.

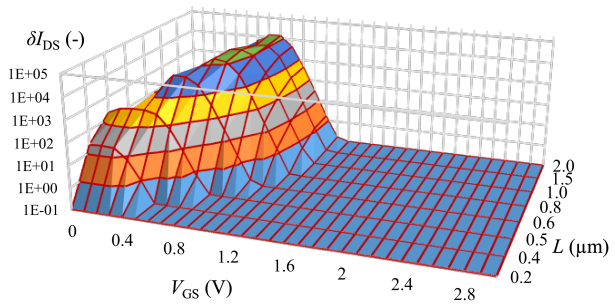


Fig. 6. The hump effect represents as drain current deviation versus gate-source voltage V_{GS} and length L of the MOS transistor

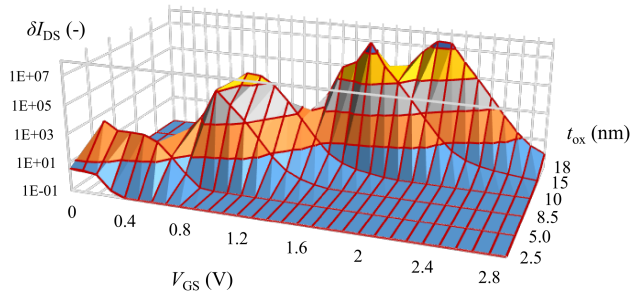


Fig. 7. The hump effect represents as drain current deviation versus gate-source voltage V_{GS} and gate oxide thickness t_{ox} of the MOS transistor

Acknowledgements

Research described in the paper was supervised by Doc. Ing. Jakovenko, Ph.D., FEE CTU in Prague and supported by the CTU SGS grant no. SGS17/188/OHK3/3T/13.

References

[1] E. A. VITTOZ. Weak inversion for ultra low-power and very low-voltage circuits. pages 129–132, Nov 2009.

[2] S. H. VOLDMAN and G. GEROSA. Mixed-voltage interface esd protection circuits for advanced microprocessors in shallow trench and locos isolation cmos technologies. *Proceedings of 1994 IEEE International Electron Devices Meeting*, pages 277–280, Dec 1994.

[3] Y. JOLY, L. LOPEZ, J. PORTAL, H. AZIZA, Y. BERT, F. JULIEN, and P. FORNARA. Impact of hump effect on mosfet mismatch in the subthreshold area for low power analog applications. *10th IEEE International Conference on Solid-State and Integrated Circuit Technology*, pages 1817–1819, Nov 2010.

[4] Silvaco TCAD Simulation Tool, DeckBuild Deck Editor Version 4.4.3.R, Synopsys Inc. 2019.

[5] P. Dimitrakis and P. Normand. Parasitic memory effects in shallow-trench-isolated nanocrystal memory devices. *Solid-State Electronics*, 51(1):147 – 158, 2007.

[6] BYOUNG-CHUL PARK, SUNG-YOUNG LEE, DONG-RYUL CHANG, KEE-IN BANG, SUNG-JUN KIM, SANG-BAE YI, and EUN-SEUNG JUNG. A fermi level controlled high voltage transistor preventing subthreshold hump. *9th International Conference on Solid-State and Integrated-Circuit Technology*, pages 172–175, Oct 2008.

[7] N. RAHMAN. Effect of threshold voltage and channel length on drain current of silicon n-mosfet. *European Scientific Journal*, 11(27):169–175, September 2015.

About Authors...

Dalibor BARRI was born in Prague, Czech Republic in 1982. He received his B.Sc. and M.Sc. degree in Electronics from the Czech Technical University (CTU), Prague, in 2005 and 2007, respectively. He worked at EMicroelectronics for five years as an analog IC front-end designer. At present time, he works at STMicroelectronics as a senior analog IC back-end physical design engineer. He is a Ph.D. student, and his topic of the thesis is to invent a novel tool for an automatic or semi-automatic layout of the analog integrated circuits or to invent and describe a layout structure with specific layout shapes.

A procedure for frequency selection of multisine impedance spectroscopy regarding higher harmonics at undersampling conditions

Alexander BLÖMEKE[1]

[1] Electrochemical Energy Conversion and Storage Systems Group, Institute for Power Electronics and Electrical Drives (ISEA), RWTH Aachen University, Jaegerstr. 17/19, 52066 Aachen, Germany

Alexander.Bloemeke@rwth-aachen.de

Abstract. *Many applications are using Impedance Spectroscopy. The Bioelectrical Impedance Analysis (BIA) and the Electrochemical Impedance Spectroscopy (EIS) are only two application examples. The author's institute uses the latter for characterizing batteries.*

Crest factor optimized multisine signals are applied in stimulation to reduce measurement time. Since the object to be analyzed is often nonlinear, the excitation signal requires special attention. If the measurement system is using undersampling the excitation signal is further restricted.

After the description of the problem, the author gives a procedure for solving this and presents an example.

Keywords

Bioelectrical Impedance Analysis, BIA, Electrochemical Impedance Spectroscopy, EIS, harmonics, multisine, Nonlinear Electrochemical Impedance Spectroscopy, NLEIS, undersampling.

1. Introduction

Using undersampling in measuring systems offers the possibility to measure frequencies higher than the Nyquist frequency. Furthermore, the required bandwidth of the measurement system can be reduced. This leads to a lower sampling frequency of the analog-to-digital converter resulting in lower costs and lower energy consumption.

Fig. 1. shows an abstract structure of impedance measurements. The excitation frequency f_e can be higher than the sampling frequency f_s, and the Low Pass Filter (LPF) is an ideal LPF with a cutoff frequency f_{LPF}. The current and voltage path are not presented separately, but for both an LPF exist and for both the same f_s is used.

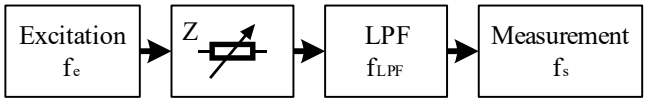


Fig. 1. Impedance measurement structure.

All spectra presented in this paper are amplitude spectra. To show the effects of harmonics and aliasing in the frequency domain, **Fig. 2.** shows a theoretical example where two frequencies are excited. The signals before the LPF are drawn with dots, the signals after filtering and undersampling with crosses. The excitation frequencies have the highest amplitude. The Device Under Test (DUT) is nonlinear which results in harmonics until the fourth Harmonic Order (HO). The second harmonic scales to the half amplitude. The third to 0.25 and the fourth to 0.1 of the original excitation amplitudes. Those values are only for demonstration. In a real system, the highest HO needs to be considered before selecting the excitation frequencies.

Based on the Nyquist-Shannon sampling theorem, all frequencies above the Nyquist frequency f_N are shifted into the 1st Nyquist zone after sampling with f_s. All Nyquist zones are having a width of $\frac{f_s}{2}$. The 1st zone starts at 0 and ends at f_N Hz. Choosing the frequencies of the excitation signal correctly allows measuring the signal itself and its harmonics in the first Nyquist zone without any overlap.

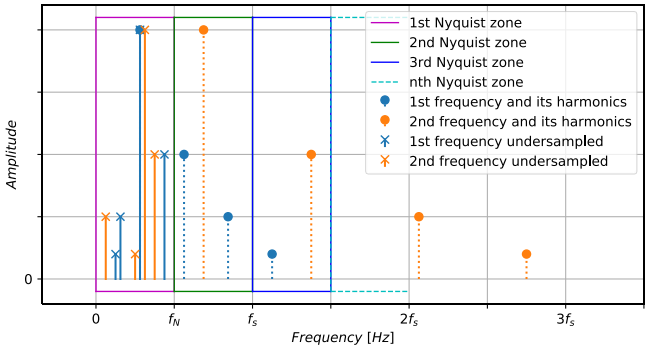


Fig. 2. Theoretical example with harmonics and aliasing.

Rees analyzes in [1] the process of selecting frequencies for a multisine signal regarding harmonics in a system that does not use undersampling. Rees shows that, for aliasing free systems, selecting prime numbers for the excitation frequencies results in overlap-free measurement results. This method is less complicated than others shown later but cannot be extended directly for undersampling systems.

In [2] and [3], an algorithm is presented by Schmitz et al., which utilizes both, the harmonics and undersampling.

However, this paper presents an alternative way of finding excitation frequencies that can be measured overlap-free. Since the problem is the same, there are some similarities to [2] and [3].

2. Procedure

First, there are some assumptions for the following analysis. The DUT is assumed to be stationary. There is a fundamental frequency $f_0 \in \mathbb{R}_{>0}$. f_0 is the greatest common divisor (GCD) of all other frequencies. Without loss of generality, in the following applies that $f_0 = 1$. This results in the fact that all frequencies are integers. After performing the procedure presented, it is possible to multiply the results with any other fundamental frequency to scale the frequency range. If selecting $f_0 = \Delta f$, where Δf is the frequency spacing of the measurement system, the density of the frequency bins is maximized.

The highest expected HO of the DUT is written as $HO_{\max}$. Furthermore, a highest measurable harmonic HO_{M} is introduced. HO_{M} defines how many harmonics the system should be able to measure. This results in a maximum frequency which can be excited:

$$f_{\max} = \left\lfloor \frac{f_{\mathrm{LPF}}}{HO_{\mathrm{M}}} \right\rfloor. \tag{1}$$

For keeping all numbers integers, the sampling frequency needs to be even, $f_{\mathrm{s}} \equiv 0\ (mod\ 2)$. This is not necessary but assumed here.

In the following graphics, it applies without loss of generality that $HO_{\max} = 4$, $HO_{\mathrm{M}} = 3$ and $\frac{f_{\mathrm{LPF}}}{f_{\mathrm{s}}} = 3.75 \Rightarrow \frac{f_{\max}}{f_{\mathrm{s}}} = 1.25$.

Fig. 3. shows the aliasing frequencies of the excitation signal and its harmonics. On the x-axis is the excitation frequency, on the y-axis the HO of each excitation frequency, and the color is showing the resulting alias frequency after lowpass filtering. As expected, the graph is symmetric around f_{N} and f_{s}. The periodicity over the frequency is increasing with the HO, since the frequency is HO-times higher than the fundamental harmonic $HO = 1$.

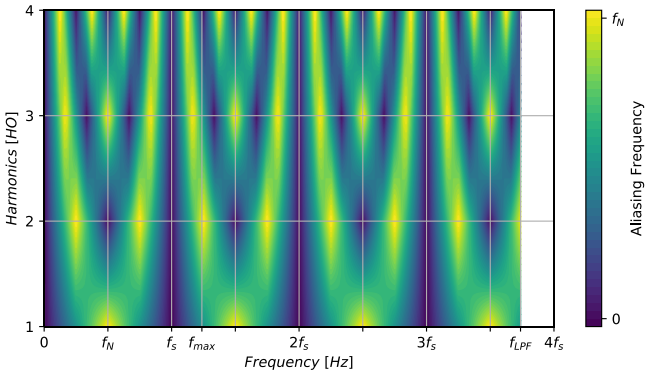


Fig. 3. Aliasing frequencies.

As a first step, only frequencies till $f_{\max}$ can be regarded, since the 3rd harmonic of frequencies higher than $f_{\max}$ is filtered out:

$$\forall f > f_{\max}: f \cdot HO_{\mathrm{M}} > f_{\mathrm{LPF}} \tag{2}$$

Based on the sampling theorem, all f_{s}-multiples are aliased to 0 Hz. To measure the constant component of the system at 0 Hz, all f_{s}-multiples are not allowed to avoid overlapping with 0 Hz. This is extendible to the harmonics. If a harmonic is equal to f_{s}, then its fundamental harmonic frequency is not allowed. The possible remaining frequencies are shown in **Fig. 4.** (white means not allowed).

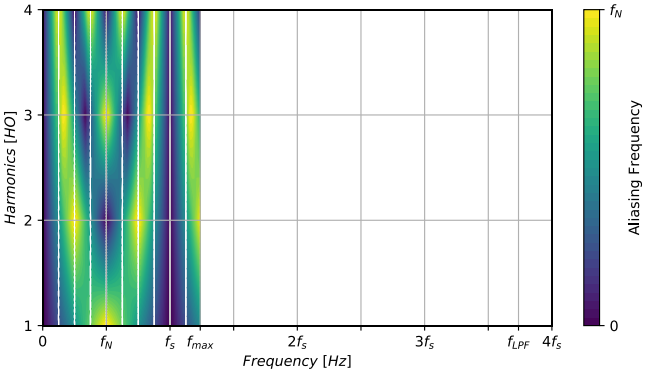


Fig. 4. Aliasing frequencies - prefiltered.

The found frequencies and their harmonics are shown in the matrix below.

$$\begin{pmatrix} f_1 \cdot HO_{\max} & \cdots & f_{\max} \cdot HO_{\max} \\ \vdots & \ddots & \vdots \\ f_1 \cdot 1 & \cdots & f_{\max} \cdot 1 \end{pmatrix}$$

After lowpass filtering "$LPF(\,)$", and sampling "$S()$" the values of the matrix are between 0 Hz and f_{N} Hz.

$$\begin{pmatrix} S\big(LPF(f_1 \cdot HO_{\max})\big) & \cdots & S\big(LPF(f_{\max} \cdot HO_{\max})\big) \\ \vdots & \ddots & \vdots \\ S\big(LPF(f_1 \cdot 1)\big) & \cdots & S\big(LPF(f_{\max} \cdot 1)\big) \end{pmatrix}$$

Two fundamental harmonics and their higher order harmonics are not overlapping after lowpass filtering and sampling if their columns in the matrix do not have any integer in common. Finding all pairs of columns that do not overlap, results in building an adjacent matrix. **Fig. 5.** shows a possible adjacent matrix.

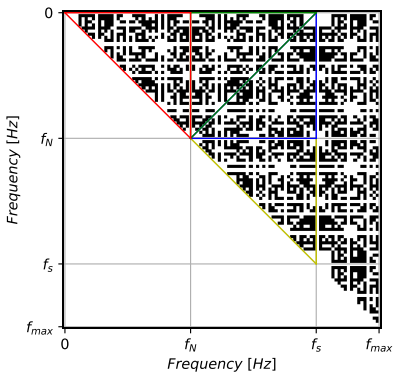


Fig. 5. Adjacent matrix (black is true, white is false).

There are lots of symmetries inside the adjacent matrix, marked with the colored triangles. Using those for further reductions might improve performance. This paper does not consider that.

From the adjacent matrix, an undirected graph is built. The nodes of the resulting graph are representing the columns in the matrix shown above. If two columns (nodes) do not have any number in common, the adjacent matrix is true (black) at this position. This results in an edge in the graph between two columns (nodes). The objective is finding frequency groups which are not overlapping, not only pairs. A clique search in the graph finds those groups. A clique in a graph defines a group of nodes which are all interlinked. In a clique of fundamental frequencies, those fundamental frequencies and their harmonics do not have any integer in common.

After finding those cliques, it depends on the application which clique is preferable. The spacing of the frequencies, like logarithmic or linear spacing, can be a selecting criterion. If the power of the signal in the time domain is limited, based on Parseval's theorem, the number of frequencies needs to be limited. This limitation keeps the Signal-to-Noise Ratio (SNR) at each frequency high enough.

After considering the best clique, the crest factor should be minimized. [4] compares different crest factor optimizations. Zappen et al. show that for different frequency spacings different optimization algorithms are performing better.

3. Example

For this example, the following parameters apply: $HO_{\max} = 4$, $HO_{\mathrm{M}} = 3$, $f_0 = 1$ Hz, $f_{\mathrm{LPF}} = 160$ Hz, and $f_{\mathrm{s}} = 42$ Hz. It follows that: $f_{\max} = \left\lfloor \frac{f_{\mathrm{LPF}}}{HO_{\mathrm{M}}} \right\rfloor = 53$ Hz and $f_{\mathrm{N}} = \frac{f_{\mathrm{s}}}{2} = 21$ Hz. The resulting graph is shown below. It consists of 39 nodes and 201 edges. The maximum clique size is 3, and there exist 30 cliques of this size.

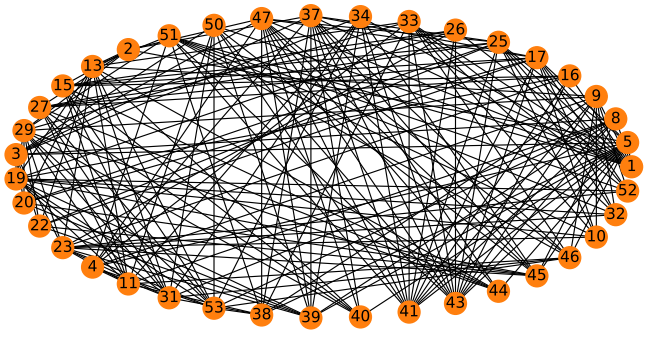


Fig. 6. Graph example.

For selecting a proper clique, in this example, the absolute differences between the frequencies in each clique are calculated. Those differences are representing the density of the frequency points. Finding the best linear frequency spread is done by considering the minimum, maximum, mean and the standard derivation of the absolute differences between the frequency points in each clique. **Fig. 7.** shows the result. Considering the maximum of the mean differences results in a linear distribution. There are two cliques with the same maximum mean. The clique with the smallest standard derivation covers a broader bandwidth. As mentioned before, this process strongly depends on the application. Finding an optimum can consider different cost functions. Considering the mean is just a simple example.

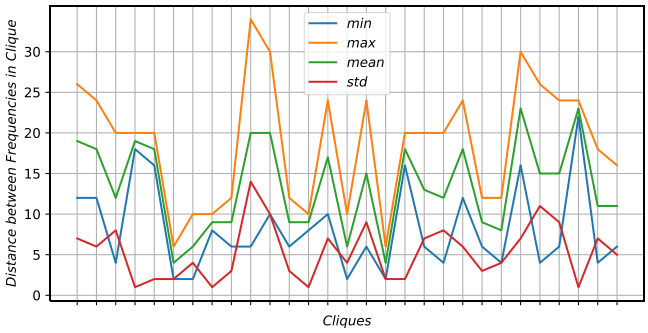


Fig. 7. Clique analysis.

The selected group consists of the frequencies: 1, 25 and 47 Hz. The matrix below contains the results. The first step is lowpass filtering, the second sampling.

$$\begin{pmatrix} 4 & 100 & 188 \\ 3 & 75 & 141 \\ 2 & 50 & 94 \\ 1 & 25 & 47 \end{pmatrix} \overset{\circlearrowleft}{\Rightarrow}_{1} \begin{pmatrix} 4 & 100 & - \\ 3 & 75 & 141 \\ 2 & 50 & 94 \\ 1 & 25 & 47 \end{pmatrix} \overset{\circlearrowleft}{\Rightarrow}_{2} \begin{pmatrix} 4 & 16 & - \\ 3 & 9 & 15 \\ 2 & 8 & 10 \\ 1 & 17 & 5 \end{pmatrix}$$

Fig. 8. shows the first and the last matrix. As in **Fig. 2.**, the harmonics having a scale of 1, 0.5, 0.25 and 0.1. The dots are representing the frequencies before lowpass filtering and sampling (first matrix). The crosses are representing the frequencies after filtering and sampling. All measured frequencies are inside the 1st Nyquist zone.

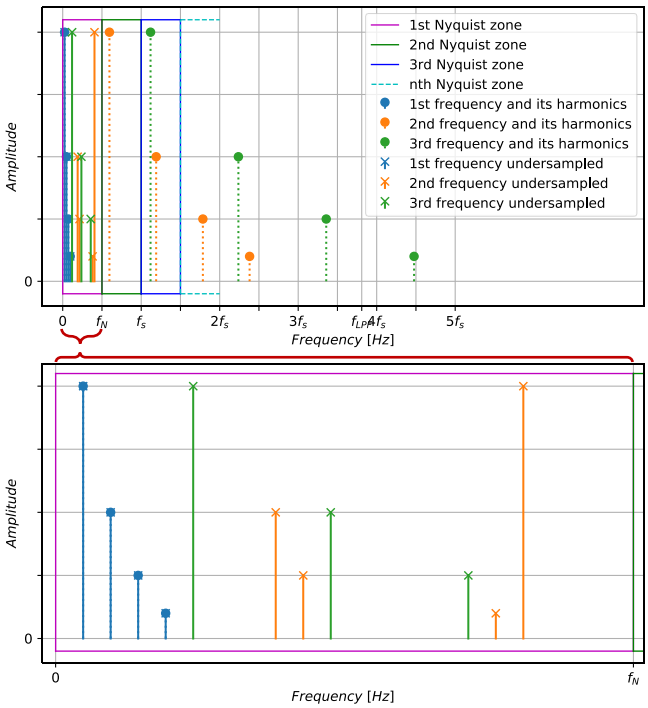


Fig. 8. Frequencies example.

4. Discussion

The presented procedure is working under the mentioned assumptions. Mapping the frequencies on integer numbers offers the possibility of applying modulo arithmetic operations and group theory. It might be the next step to propose a general mathematical description for the group of frequencies and their relation.

Modifying the sampling frequency f_s by small values changes the maximum clique size and the number of cliques in a non-obvious behavior. With a mathematical description, this might be predictable and therefore be considered at system planning. **Tab. 1.** shows this behavior. All other parameters are kept the same as in the example above.

f_s	Maximum clique size	Number of cliques of maximum size
48	3	272
46	4	690
44	5	210
42	3	30
40	2	116
38	3	282

Tab. 1. Number of cliques and size over f_s.

5. Conclusion

A general procedure of finding frequencies for multisine signals at undersampling is described. As presented, the problem is mapped on an undirected graph and solved by taking the maximum cliques. Describing the problem with a mathematical group theory might improve the algorithm.

Summarizing the procedure:

- Calculate $f_{\max} = \left\lfloor \frac{f_{\mathrm{LPF}}}{HO_{\mathrm{M}}} \right\rfloor$
- Create a matrix, that contains all numbers from $f_0 = 1$ till $f_{\max}$ and its harmonics

$$\begin{pmatrix} HO_{\max} \\ \dots \\ 1 \end{pmatrix} \cdot (f_0 = 1 \quad \dots \quad f_{\max}) = \begin{pmatrix} f_1 \cdot HO_{\max} & \cdots & f_{\max} \cdot HO_{\max} \\ \vdots & \ddots & \vdots \\ f_1 \cdot 1 & \cdots & f_{\max} \cdot 1 \end{pmatrix}$$

- Delete all columns that contain a multiple of f_s
- Calculate the frequencies after lowpass filtering "LPF" and sampling "S"

$$\begin{pmatrix} S\big(LPF(f_1 \cdot HO_{\max})\big) & \cdots & S\big(LPF(f_{\max} \cdot HO_{\max})\big) \\ \vdots & \ddots & \vdots \\ S\big(LPF(f_1 \cdot 1)\big) & \cdots & S\big(LPF(f_{\max} \cdot 1)\big) \end{pmatrix}$$

- Build an adjacent matrix, representing if two columns are sharing a number or not
- Create an undirected graph form this adjacent matrix
- Make a maximum clique search

Based on the application continue with:

- Selecting the best maximum clique
- Minimize the crest factor of the multisine signal created with the selected frequencies

Acknowledgments

Hendrik Zappen supervised the research described in the paper. Thanks to Dirk Uwe Sauer for making the author's master thesis possible.

References

[1] DAVID REES, *Digital Processing of System Responses*. (Doctoral dissertation, Polytechnic of Wales). Treforest (UK), 1976.

[2] MICHAEL J. SCHMITZ, ROGER A. GREEN, *Optimization of Multisine Excitations for Receiver Undersampling*. 2010 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, Dallas (USA), 2010.

[3] MICHAEL J. SCHMITZ, *Multisine Excitation Design to Increase the Efficiency of System Identification Signal Generation and Analysis*. (Doctoral dissertation, North Dakota State University). North Dakota (USA), 2012.

[4] HENDRIK ZAPPEN, FLORIAN RINGBECK, DIRK UWE SAUER, *Application of Time-Resolved Multi-Sine Impedance Spectroscopy for Lithium-Ion Battery Characterization*. Batteries 4(4):64. 2018.

About Authors...

Alexander BLÖMEKE has studied electrical engineering at the RWTH Aachen University (Germany), with focus on information and communication technology. He will graduate in March 2019 and is currently writing his master thesis.

Selectivity enhancement of target volatile organic compounds using sensor array and temperature modulation

Ali AHMAD[1]

1 Dept. of Microelectronics, Faculty of Electrical Engineering, CTU in Prague, Technicka 2, 16627 Prague, Czech Republic

ahmadali@fel.cvut.cz

Abstract. *In this paper, we report the performance of a volatile organic compounds identification system based on a commercial metal-oxide gas sensor array. Analyses of five volatile organic compounds (heptane, pentane, isooctane, isopropyl alcohol, and tetrachloroethylene) Have been performed with these unselective sensors. The device consists of three MOX commercial sensors (AS-MLN, AS-MLK, AS-MLC), The behavior of these sensors was measured at different ranges of temperature. Principal component analysis (PCA) was applied in order to identify the target gases. We obtained a high classification of the five VOCs by temperature modulation. The final aim of this research is to improve the selectivity of commercial sensors based on metal oxide to classify different VOCs.*

Keywords

Gas sensor array, metal oxides, selectivity enhancement, temperature modulation, principal component analysis.

1. Introduction

The measurements of volatile organic compounds (VOCs) are becoming more important due to stringent environmental regulations and increasing health concerns [1], also Human breath includes many volatile organic compounds (VOCs) that can be used as biomarkers for diseases. Therefore, breath-monitoring diagnostic methods are desirable as diagnostic tools because they are fast and non-invasive [5]. Consequently, many researchers have attempted to develop breath-monitoring systems [6–7]. Metal oxide (MOX) sensors are well known as multifunction materials and employing MOS in detecting VOCs is one of the most studied areas [3], [4]. The advantages of metal oxide sensors in comparison with sensors based on other materials are well known like low costs, short response time and versatility. Currently, these sensors are sufficiently sensitive for most applications [8–11]. However, the use of them is limited due to their lack of selectivity. which has stimulated researchers to look for different strategies to overcome this drawback, Recent works have shown that temperature modulation of metal oxide sensors improves selectivity [2],[13–19]. The kinetics of adsorption and desorption are modified when the temperature of the sensor is modulated, and the sensors have specific response patterns to different gases [12]. Usually, an electronic nose measures the stationary response of the sensors working at one constant temperature, but because the response of a sensor depends on its working temperature, the combination of the standard approach and a temperature stepping increases the number of signatures specific to each gas without increasing the number of sensors. The improvement of selectivity is investigated and reported in this paper by using a portable device consists of three metal oxide commercial gas sensors (AS-MLN, AS-MLK, AS-MLC) with modulation of the heating voltage to classify and identify different volatile organic compounds. We performed principal component analysis (PCA) to analyze the data extracted by exposing different VOCs during the temperature change.

2. Experimental:

The analyzed data sets were gathered from the portable gas detector Fig.1 which was made and equipped with three commercial conductivity gas sensors (Applied Sensor) with different MOX thin sensing films: AS-MLN sensitive mainly on NO_2, AS-MLC sensitive mainly on CO, and AS-MLK sensitive mainly on CH_4 [20].

Fig.1 Portable gas detector [20]

These sensors require low heating power 30 - 45 mW. Due to their low thermal capacity, these sensors react quickly to

the temperature changes of the heating plate. The sensor with dimensions 2 x 2 x 0.38 mm is placed in the metallic case TO-39-4. Two pins relate to the gas sensor layer for the measurement of its resistivity changes, two other pins are for the heating element (Fig.2). The sensor AS-MLN based on the WO_3 film has the sensitivity of 0.1 ppm – 2 ppm of NO_2 at the temperature of 270°C. The sensor AS-MLK based on the SnO_2 film has the sensitivity of 0.01% to 4%. of methane at the temperature 320°C. The sensor AS-MLC based on the SnO_2 film has the sensitivity 0.5ppm to 500ppm of CO at the temperature 270°C. The detector measures the sensor resistivities and the heating elements resistivities for the temperature regulation. The measurement was made at the constant flow of prepared gas mixtures with different concentrations and temperatures during several cycles. The flow meters and regulators kept the specified gas concentrations of specific gases in the carrier gas. The purging of the sensing chamber has been realized in each cycle.

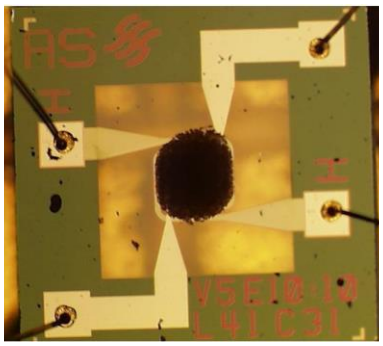
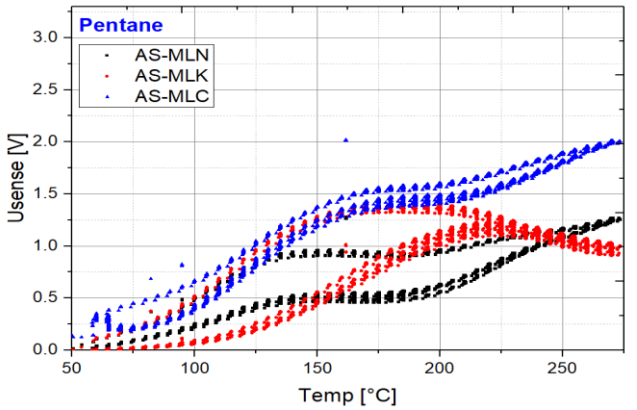
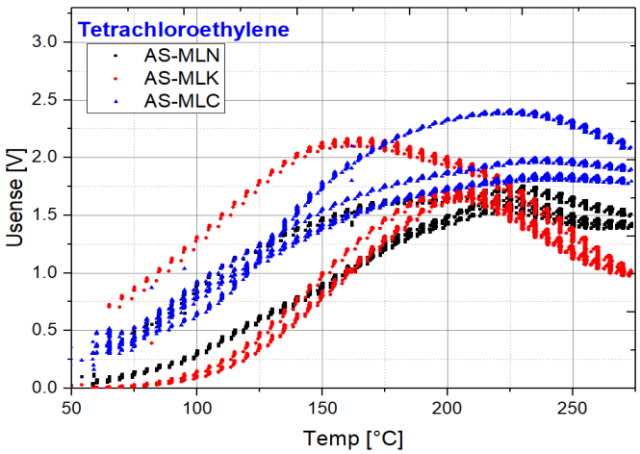
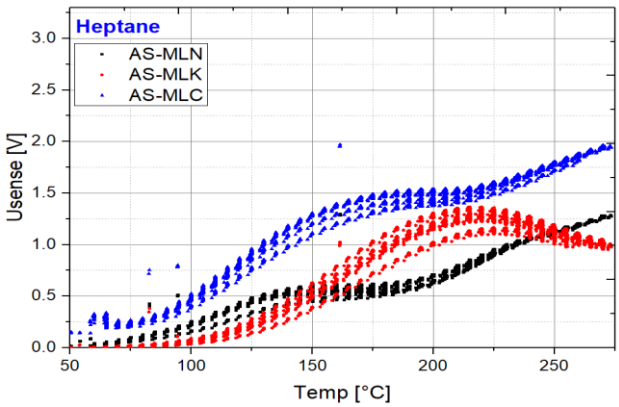
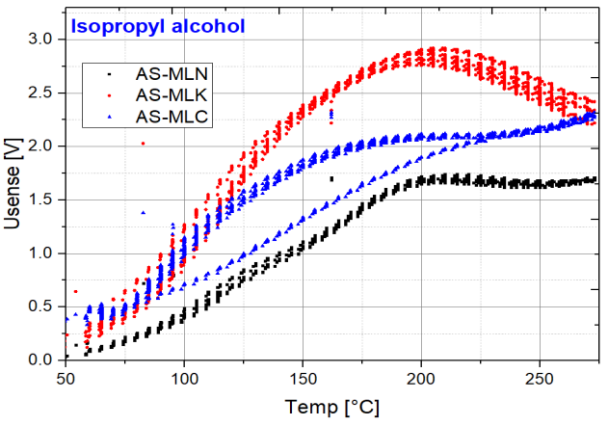


Fig. 2 Sensor AS-MLX (The case and detailed view.) [20]









3. Results and discussion:

3.1 Characterization of gas sensors:

The sensor Characterization was initially obtained by exposing five volatile organic compounds (VOCs) to an array of sensors in the temperature range of (50 –260) ° C, Fig.3 presents the response of the three commercial metal oxide sensors AS-MLN, AS-MLK, AS-MLC to five volatile organic compounds, the result showed that the sensors have very good sensitivity to the target gases, sufficiently stable and have good reproducibility which is very important for most of the applications .

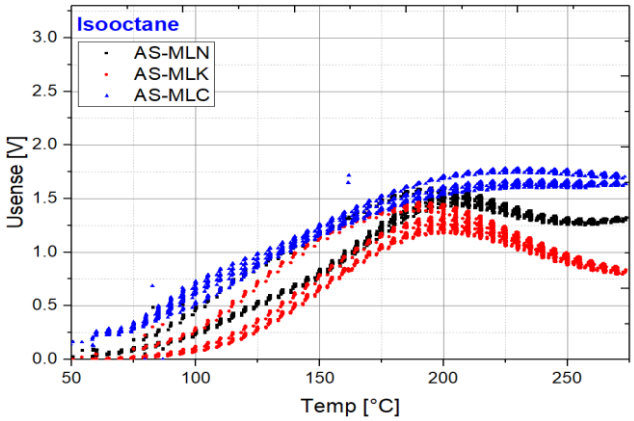


Fig. 3 the response of the three sensors AS-MLN, AS-MLK, AS-MLC to five volatile organic compounds in the temperature range of (50 –260) ° C

3.2 Discrimination of the target gases:

In order to obtain more information about the system, PCA was performed to discriminate among five compounds.

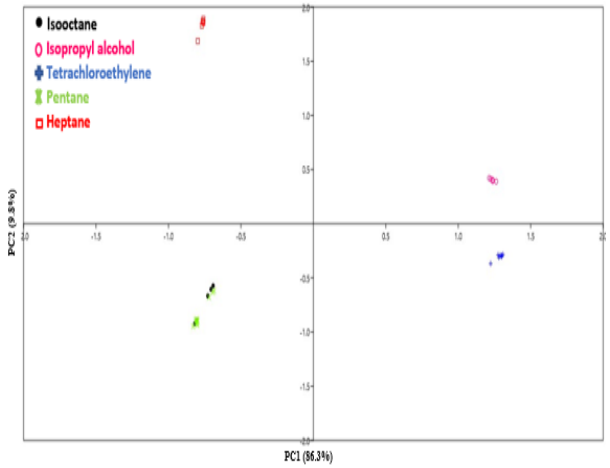


Fig.4 Discrimination of five compounds by three MOS commercial sensors in the temperature range (60 to 120) °C

Fig.4 presents the result of applying PCA on the dataset which extracted during increasing the temperature in the range of (60 to 120) °C, The first principal component describes 86.3% of data variation and PC2 describes 9.8% of data variation which is neglected in comparison with PC1,so we will focus in our comparison on the variation of PC1, Fig.4 shows an overlap between heptane, pentane, and isooctane which they have approximately the same influence on PC1 and also overlap between Isopropyl alcohol and Tetrachloroethylene. The system has poor classification with a low range of temperature. However, a high classification of the five compounds in clearly separated clusters is obtained in Fig.5 with a temperature range of (160 to 220) °C The experimental results of this work points out that the capability for gas identification is hugely improved for commercial sensors by temperature modulation.

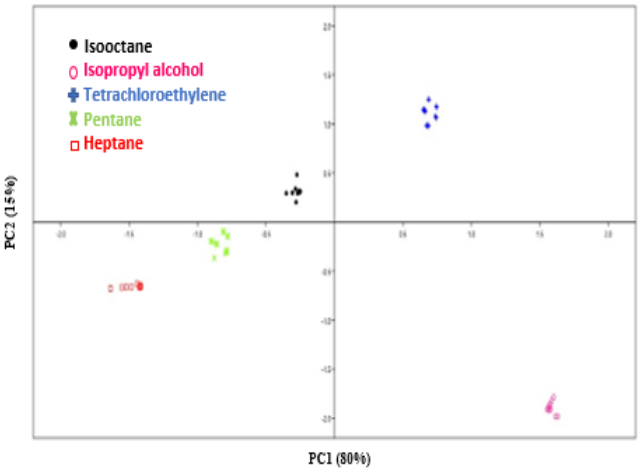


Fig.5 Discrimination of five compounds by three MOS commercial sensors In the temperature range (160 to 220) °C

4. Conclusions:

In this work, we studied the ability to carry out gas analyses with a metal-oxide sensor array, the advantages of these sensors are their high commercial availability and low cost. PCA was performed as a pattern recognition method in order to identify different compounds (heptane, pentane, isooctane, isopropyl alcohol, and tetrachloroethylene). The results pointed out that sensors have very good classification in a higher range of temperature, also pointed out to the importance of a sensor temperature programming procedure and confirm that different temperatures give more information about the chemical response.In spite of the low selectivity of MOS sensors, the improvement of gas identification is realizable with cheap sensors and a simple temperature programming method.

Acknowledgments

The research described in the paper was supervised by doc. RNDr. Jan Voves, CSc., FEL CTU in Prague.

References

[1] TOSHIO ITOH *, TAKAFUMI AKAMATSU, AKIHIRO TSURUTA,AND WOOSUCK SHIN selective Detection of Target Volatile Organic Compounds in Contaminated Humid Air Using a Sensor Array with Principal Component Analysis.

[2] A.FORT *, M.GREGORKIEWITZ , N.MACHETTI , S.ROCCHI,B.SERRANO , L.TONDI , N.ULIVIERI Selectivity

enhancement of SnO sensors by means of operating temperature modulation.

[3] J. HUANG, Q. WAN Gas sensors based on semiconducting metal oxide one-dimensional nanostructures Sensors, 12 (2009), pp. 9903-9924.

[4] N.H. AL-HARDAN, M.J. ABDULLAH, A. ABDUL azizsensing mechanism of hydrogen gas sensor based on RF sputtered ZnO thin films Int J Hydrogen Energy, 35 (2010), pp. 4428-4434.

[5] AMANN, A.; SPANEL, P.; SMITH, D. Breath analysis: The approach towards clinical applications. MiniRev. Med. Chem. 2007, 7, 115–129. [CrossRef] [PubMed].

[6] PENG, G.; TISCH, U.; ADAMS, O.; HAKIM, M.; SHEHADA, N.; BROZA, Y.Y.; BILLAN, S.; BORTNYAK, R.A.; KUTEN, A.; HAICK, H. Diagnosing lung cancer in exhaled breath using gold nanoparticles. Nat. Nanotech. 2009, 4, 669–673.

[7] MAZZONE, P.J.; HAMMEL, J.; DWEIK, R.; NA, J.; CZICH, C.; LASKOWSKI, D.; MEKHAIL, T. Diagnosis of lung cancer by the analysis of exhaled breath with a colorimetric sensor array. Thorax 2007, 62, 565–568.

[8] C. PIJOLAT, C. PUPIER, M. SAUVAN, G. TOURNIER, R. LALAUZE, Gas detection for automotive pollution control, Sens. Actuators B 59 (1999) 195 202.

[9] M. PENZA, G. CASSANO, F. TORTORELLA, G. ZACCARIA, Classification of food, beverages and perfumes by WO3 thin film sensors array and pattern recognition techniques, Sens. Actuators B 73 (2001) 76–87.

[10] K. AGUIR, C. LEMIRE, D. LOLLMAN, Electrical properties of reactively sputtered WO3 thin films as ozone gas sensor, Sens. Actuators B 84 (2002) 1–5.

[11] D.S. LEE, Y.T. KIM, J.S. HUH, D.D. LEE, Fabrication and characteristics of SnO2 gas sensor array for volatile organic compounds recognition, Thin Solid Films 416 (2002) 271–278.

[12] KIEU AN NGO, PASCAL LAUQUE ∗, KHALIFA AGUIR .High performance of a gas identification system using sensor array and temperature modulation.

[13] A.P. LEE, B.J. REEDY, Temperature modulation in semiconductor gas sensing, Sens. Actuators B 60 (1999) 35–42.

[14] A. ORTEGA, S. MARCO, A. PERERA, T. SUNDIC, A. PARDO, J. SAMITIER, An intelligent detector based on temperature modulation of a gas sensor with a digital signal processor, Sens. Actuators B 78 (2001) 32.

[15] A. FORT, M. GREGORKIEWITZ, N. MACHETTI, S. ROCCHI, B. SERRANO, L. TONDI, N. ULIVIERI, V. VIGNOLI, G. FAGLIA, E. COMINI, Selectivity enhancement of SnO2 sensors by means of operating temperature modulation, Thin Solid Films 418 (2002) 2–8.

[16] N.J. CHOI, C.H. SHIM, K.D. SONG, D.S. LEE, J.S. HUH, D.D. LEE, Classification of workplace gases using temperature modulation of two SnO2 sensing films on substrate, Sens. Actuators B 86 (2002) 251–258.

[17] R. IONESCU, E. LLOBET, WAVELET transform-based fast feature extraction from temperature modulated semiconductor gas sensors, Sens. Actuators B 81 (2002) 289–295.

[18] A. GRAMM, A. SCHUTZE, High performance solvent vapor identification with ¨ a two sensor array using temperature cycling and pattern classification, Sens. Actuators B 95 (2003) 58–65.

[19] R. IONESCU, A. HOEL, C.G. GRANQVIST, E. LLOBET, P. HESZLER, Ethanol and H2S gas detection in air and in reducing and oxidizing ambience: application of pattern recognition to analyses the output from temperature-modulated.

[20] J. NEMAZAL: Portable gas detection monitor, master thesis, CTU Prague, 2015.

About Authors...

Ali AHMAD was born in Latakia in 1993. He graduated in communication and electronic engineering at the Tishreen University in Latakia, now he is a PhD student in FEL, CTU in Prague, Dept. of Microelectronics.

High Performance Ultra-low Quiescent Current LDO Regulator

Lukas Buryanec[1,2]

[1]STMicroelectronics, Pobrezni 3, Prague, Czech Republic
[2]Dept. of Microelectronics, Czech Technical University, Technicka 2, 166 27 Prague, Czech Republic

lukas.buryanec@st.com, buryaluk@fel.cvut.cz

Abstract. *This paper deals with the design of a high-performance ultra-low quiescent current LDO voltage regulator in BCD technology of STMicroelectronics company. Efficiency is one of the most important aspects of contemporary integrated circuits. Therefore, besides the ultra-low quiescent current around 1 microampere under no load current condition, the emphasis has been put on ultra-low drop-out voltage as well. Although there are available ultra-low quiescent current LDOs on the market with a quiescent current in units of microamperes, some of their parameters are not usually sufficiently good, for example, power supply rejection ratio (PSRR) and transient response behavior. The mentioned ultra-low drop-out voltage is not generally their boast feature either. The proposed LDO represents a unique device comparing to commercial products regarding the mentioned parameters. Another key feature of this device is the ability to function even with as low input voltage as 0.8 volts. This is allowed by using an NMOS power transistor while the vast majority of commercial LDOs uses PMOS transistors which normally do not allow that low input voltage.*

Keywords

Poster2019, IC design, power management

1. Introduction

Low drop-out regulators (LDOs) are linear voltage regulators capable of maintaining the input voltage even with its very low difference from the regulated output voltage. This difference (drop-out voltage) can be as low as tens to hundreds of millivolts. Linear regulators functionality is based on a principle of dissipating energy across a pass element. Thanks to the very low drop-out voltage across the pass element, very high efficiency can be achieved. Nowadays, battery-powered electronic devices such as cell phones are indivisible parts of everyday life. Thus, efficiency is one of the key features for a long battery endurance of these devices. Besides the low drop-out voltage, there is another factor in terms of efficiency and it is a low quiescent current which is necessary especially when the LDO is operated under no load current conditions.

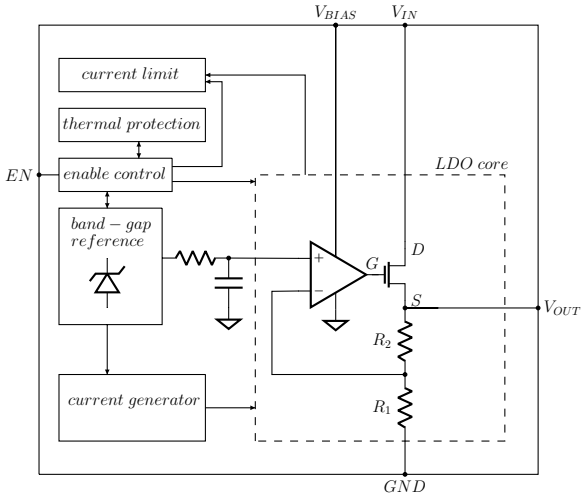


Fig. 1. Simplified block schematic of the designed LDO regulator

2. Complications of low-quiescent current design

Achieving an ultra-low quiescent current of LDOs brings complications and difficulties during the design. The main obstacles can be following:

1. Minimizing the number of current branches restricts the freedom of the design.
2. Attaining the very low drop-out voltage causes using a large power transistor. The large area of MOS transistors also represents a large gate capacitance. For that reason, driving the transistor is difficult with a low current.
3. The large gate capacitance makes an internal pole of the LDO. Thus, problems regarding stability can occur. Pushing this poles outside of the open-loop bandwidth without a sufficiently large current and other frequency compensation techniques can represent a key challenge.

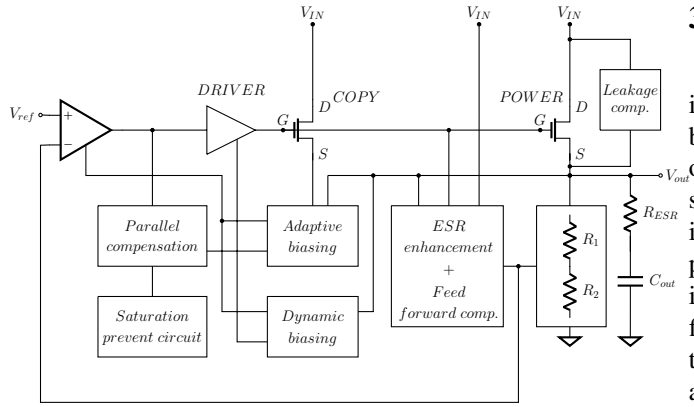


Fig. 2. Simplified block schematic of the designed LDO core

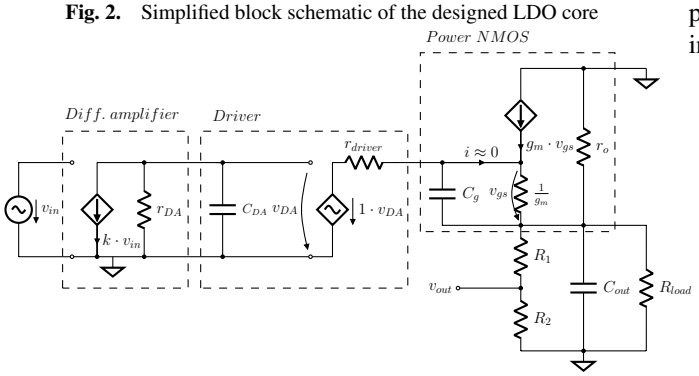


Fig. 3. Simplified small-signal model of the designed LDO regulator for determining its open-loop gain and poles

4. For high PSRR at higher frequencies, the large bandwidth is one of the main contributors. For its achieving, the large current is needed as well.

5. For sufficient transient response behavior, the large bandwidth is needed again and capacitances in the system have to be charged/discharged fast enough. The problem with the large gate capacitance and low current occurs again.

3. Structure of the proposed LDO

The block schematic of the proposed LDO is depicted in Fig. 1. The circuit contains, besides the LDO core, a band-gap voltage reference, a current reference generator, enable logic, thermal protection, and a current limiting circuit. Since the LDO uses the NMOS power transistor, for the ultra-low drop-out capability, it is necessary to provide a sufficiently high voltage headroom for V_{GS} gate-source voltage of the power NMOS. In this case, separate biasing voltage V_{BIAS} in the range of 2.5 to 5.5 volts and input voltage V_{IN} have been preferred over usage of a charge pump. This voltage supplies the opamp and other circuitry in order to control the gate of the power NMOS as depicted in Fig. 1. This paper is focused mainly on the LDO core of which structure with main function blocks is depicted in Fig. 2.

3.1. Stability and compensation overview

The biggest challenge in designing any LDO can be, in many times, assuring stability under all conditions. For better understating and investigation of stability and the open-loop transfer function, there is a simplified small-signal model of the designed LDO for determining its poles in Fig.3. The main feedback loop contains a differential amplifier, a driver of the power transistor, the power NMOS itself and a voltage feedback divider. The differential amplifier practically represents the only one gain stage in the system since the structure of the proposed driver acts as a voltage follower as well as the connection of the power NMOS. From this model, it can be seen that there are three "main" poles formed by resistances and capacitances at each block interface. The poles can be expressed as:

$$p_1 = -\frac{1}{r_{DA}C_{DA}}, \tag{1}$$

$$p_2 = -\frac{1}{r_{driver}C_g}, \tag{2}$$

$$p_3 = -\frac{1}{r_{out}C_{out}}, \tag{3}$$

where for the output resistance r_{out} we can write

$$r_{out} = \frac{1}{g_m} \parallel r_o \parallel (R_1 + R_2) \parallel R_{load} \approx \frac{1}{g_m}, \tag{4}$$

this simplifying may be done considering that the resistors of the feedback divider are quite large, usually, in the order of megaohms, the output resistance r_o of the NMOS is large as well and its transconductance g_m is high enough due to its high $\frac{W}{L}$ ratio. Of course, this model is simplified, in reality, there will be other high frequency poles and zeros in the system.

Since there are the three mentioned poles, the system may be unstable. The chosen compensation strategy has been following: pushing the driver pole p_2 to frequencies as far as possible, pushing pole p_1 to lower frequencies by active parallel compensation which also assures output pole tracking and inserting of a "variable" zero into the system according to the output pole p_3 movement with the load current. Besides the main compensation, there are other "supportive" circuits based on feed-forwarding compensation technique and enhancement of equivalent series resistance (ESR) of the output capacitor, this resistance also inserts a zero into the system.

3.2. Block overview

A) *Power NMOS*: The drop-out voltage is determined by the resistance of the power NMOS in fully on-state - $R_{DS_{on}}$. From the standard equation of NMOS transistors

operated in linear region it follows that the main factor, besides others, for achieving low $R_{DS_{on}}$ is large $\frac{W}{L}$ ratio of the NMOS. In our case, it results in a very large size, thus large gate capacitance.

B) *Driver*: The driver is designed the way that it acts as a voltage follower. The biggest challenge during the design of the driver has been assuring its output resistance as low as possible to push the pole p_2 at the driver/power NMOS interface outside of the bandwidth and supplying a sufficient amount of current necessary for a fast slew rate and fast transient response behavior. Both goals are in a breach of the low-quiescent current. For that reason, adaptive biasing has been implemented. The driver has its own non-linear adaptive biasing - the quiescent current of the driver is a non-linear function of the load current. The adaptive biasing does not represent such a large decreasing of efficiency since at high loads, there is not a need for the quiescent current in the order of nanoamperes. The non-linear function is implemented mainly for purposes at the middle load current range when it is most difficult to move the driver pole outside of the bandwidth. Also, a sufficient current through the driver at this load current range is necessary for the good slew rate and reduction of undershoots at the output voltage during very large transient load current steps. A part of the driver is an implemented mechanism how to prevent a classic problem in LDOs - steep increase of the quiescent current when entering the drop-out mode - usually, this value of the current is not generally presented in datasheets of commercial products. The implemented mechanism clamps a current through the driver at a certain level, thus improving efficiency in the drop out mode. A good example of a drop-out mode current clamp for PMOS regulators can be found in [1].

C) *Differential amplifier*: Since the driver acts as a voltage follower and the output voltage of the LDO is adjustable in the range from 0.8 to 4 volts, there is a need for a large output swing of the differential amplifier. Therefore, the structure is based on a so-called symmetrical operational transconductance amplifier. To improve the overall open-loop gain and also PSRR, several low-threshold cascoding current mirrors have been implemented into the structure. To increase the open-loop bandwidth, thus increase PSRR, and the slew rate, adaptive biasing has been used again. Now, it is ensured by the precise adaptive biasing block which we be described later.

D) *Active parallel compensation*: As it has been already said, the active parallel compensation is the main compensation technique and it uses the fact that the output pole p_3 of the LDO moves with the load current. The simplified principle is based on the precise adaptive biasing current which is used to control a "variable" zero formed by a transistor operated in linear region. Several complications occurred during the design, for instance, leakage current at high temperatures caused assuring a sufficient phase margin very problematic namely at light loads when the adaptive biasing current may be in the same order as the leakage current. Thus, a leakage current compensation for this block has been implemented as well.

E) *Saturation preventing for parallel compensation*: During very large transient steps, under some conditions, it can happen that the zero making transistor working in linear region can potentially enter saturation. This can occur as ringing at the output voltage. This block ensures that the V_{DS} voltage of the transistor does not cross the minimal saturation voltage level. It is based on monitoring the V_{DS}. If this voltage is about to cross a certain level, the circuit activates and clamps the V_{DS} at this value.

F) *Precise adaptive biasing*: To provide precise mirroring of the load current by the copy NMOS transistor, it is necessary to assure that its both V_{GS} and V_{DS} are the same as in the power NMOS - this is the purpose of this block followed by current mirroring to other blocks.

G) *Dynamic biasing*: During very large load transients steps, the adaptive biasing current of the differential amplifier and the driver may not be sufficiently large for the fast slew rate - fast recharging of capacitances, mainly the compensation capacitor and the gate capacitance of the power NMOS. When there is about to happen an undershoot at the output voltage during transients, the dynamic biasing block boosts a large amount of current to the differential amplifier and to the driver as well. This significantly improves the transient response and reduces undershoots. At the steady state, this block does not boost practically any current. This also represents the biggest challenge - assuring that no current is boosted under every condition in the steady state, mainly during Monte Carlo simulations.

I) *ESR enhancement and feed-forward compensation*: The LDO may be operated with output ceramic capacitors with a very low ESR in units of milliohms. That low values may not have any impact on improving the phase margin. This block "virtually" enhances the ESR and improves the phase margin. It is based on making another high frequency regulation loop path which is separated from the DC regulation loop by a high pass filter. This circuit also allows using a feed forward compensation technique even when the output voltage of the LDO is set to the same level as the output voltage of the band-gap reference - in that case, the feedback divider is not connected to the feedback signal. Therefore, the standard feed-forward compensation cannot be used in that case.

J) *Leakage compensation*: As it has been said, the $\frac{W}{L}$ ratio of the power NMOS is very large. This also means that the leakage current of the NMOS at high temperatures is not negligible and cannot be left ignored. If so, the voltage regulation at light loads may not apply - the power NMOS may not be turned-off sufficiently. The leakage compensation is based on a generation of a proportional leakage current and subtracting it from the power transistor path branch - the voltage regulation then works normally at high temperatures at light loads.

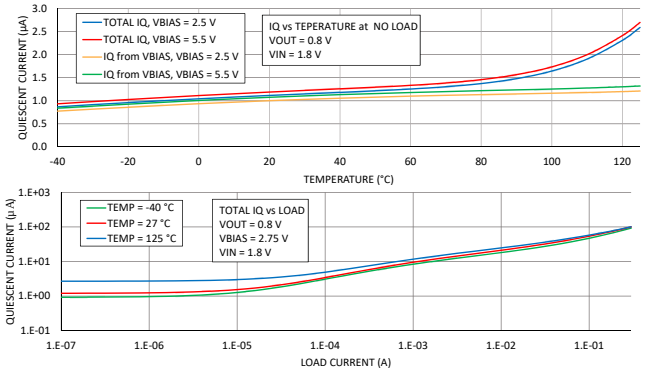


Fig. 4. Quiescent current versus temperature and load current

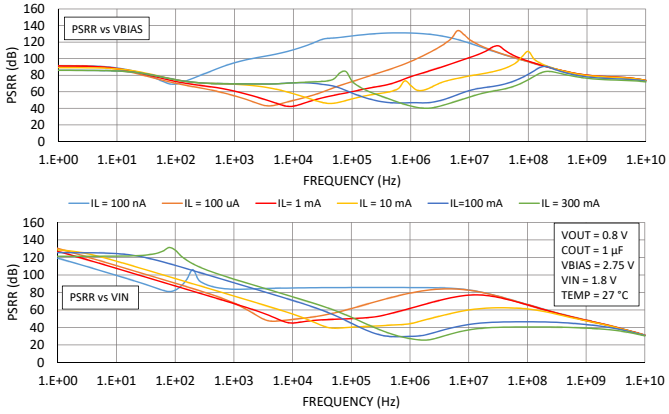


Fig. 5. PSRR versus frequency

4. Conclusion

A high performance ultra-low quiescent LDO regulator has been designed for purposes of STMicroelectronics company in BCD technology. According PC simulations using Eldo simulator the results are following: The quiescent current of the complete chip is only around 1.2 μA at no load and the room temperature. Dependence of the quiescent current on temperature and the load is depicted in Fig. 4. The increase at high temperatures is caused by the leakage current of the large power NMOS transistor. An extraordinary low drop-out voltage (vs. V_{IN}) - around 30 mV at 300 mA load current has been achieved. The controlling voltage V_{BIAS} operating value is expected to be 1.5 V or more above the output voltage. This value is standard in similar category of commercial products with different power and controlling

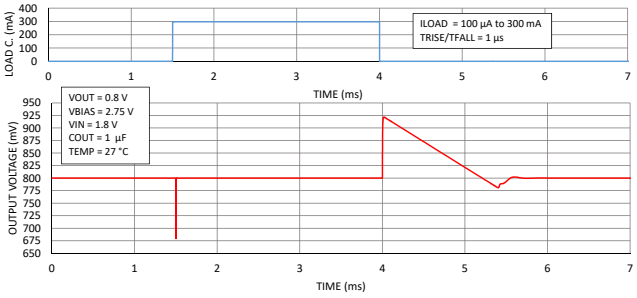


Fig. 6. Load transient response

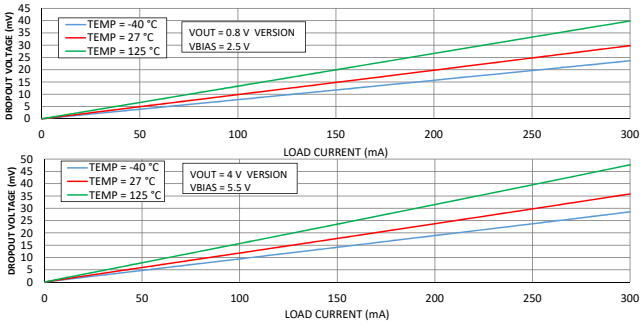


Fig. 7. Dropout voltage versus load current

supply voltages. The dropout voltage versus V_{BIAS} is in the worst case around 1.1 V for the maximal load current.

An effort has been spent on reaching good PSRR and transient response behavior which are traditionally in a breach of low quiescent current. PSRR is approximately around 70 dB at 100 Hz versus V_{BIAS} and around 80 dB versus V_{IN} which is shown in Fig. 5. Improving PSRR at high frequencies can be potentially based on [9]. The dynamic biasing has been implemented for transient response improvement - load transient response is depicted in Fig. 6. The ability of achieving the value of the mentioned drop-out voltage even at the minimal output voltage of 0.8 V is unique as well - as can be seen in Fig. 7. The output voltage is adjustable by 50 mV according to a customer with the maximal value of 4 V.

Other important parameters are on a high-performance level as well, for instance, the line and load regulation. An offer of ultra-low quiescent current LDOs on the market is restricted. The comparison of "difference making" parameters in the ultra-low quiescent category of the proposed LDO with other commercial products of respected companies - competitors of STMicroelectronics on the market is depicted in Table 1. It has to be said that it is difficult to compare the products because the given parameters in datasheets may not be measured under the completely same conditions for each LDO. Nevertheless, it can be seen that the achieved parameters represent a unique device.

Acknowledgements

Research was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS17/188/OHK3/3T/13. Also, I would like to thank my supervisor Jiri Jakovenko (Dept. of Microelectronics, Czech Technical University, Technicka 2, 166 27 Prague, Czech Republic) and my colleague Sandor Petenyi (STMicroelectronics, Pobrezni 3, Prague, Czech Republic) for their advices and guidance during designing of this device.

TYPE:	this LDO	TPS7B82Q1	TLV704	TPS783	NCP718	NCP583	NCP170	NCP4624
MANUFACTURER:	ST	TI [2]	TI [3]	TI [4]	OnSemi [5]	OnSemi [6]	OnSemi [7]	OnSemi [8]
I_Q (μA) @ $I_{LOAD} = 0, T \approx 27$ °C	1.2	1.9	3.2	0.5	4	1	0.5	2
$V_{IN_{min}}$ (V)	0.8 + drop	3 + drop	2.5	2.2	2.5	1.7	2.2	2.5
$V_{OUT_{min}}$ (V)	0.8	3	1.2	1.8	1.2	1.5	1.2	1.2
$I_{LOAD_{max}}$ (mA)	300	300	150	150	300	150	150	150
$V_{DROPOUT}$ (mV) (vs. V_{IN}) @ $I_{LOAD_{max}}, T \approx 27$ °C	30	600	1600	130	480	600	350	1680
$PSRR$ (dB) @ 100 Hz, $I_{LOAD} \approx 10$ mA, $T \approx 27$ °C	80 (vs. V_{IN}) 70 (vs. V_{BIAS})	60	50	20	70	50	65	40

Tab. 1. Comparison with commercial products

References

[1] PETENYI, S. *Voltage regulator with dropout detector and bias current limiter and associated methods*. US Patent US9645594, 2017.

[2] Texas Instruments. *TP37B82-Q1 [online]*. [cit. 2018-01-01]. Available from: http://www.ti.com/lit/ds/symlink/tps7b82-q1.pdf

[3] Texas Instruments. *TLV704 [online]*. [cit. 2018-01-01]. Available from: http://www.ti.com/lit/ds/symlink/tlv704.pdf

[4] Texas Instruments. *TPS783 [online]*. [cit. 2018-01-01]. Available from: http://www.ti.com/lit/ds/symlink/tps783.pdf

[5] OnSemiconductor. *NCP718 [online]*. [cit. 2018-01-01]. Available from: http://www.onsemi.com/pub/Collateral/NCP718-D.PDF

[6] OnSemiconductor. *NCP583 [online]*. [cit. 2018-01-01]. Available from: http://www.onsemi.com/pub/Collateral/NCP583-D.PDF

[7] OnSemiconductor. *NCP170 [online]*. [cit. 2018-01-01]. Available from: http://www.onsemi.com/pub/Collateral/NCP170-D.PDF

[8] OnSemiconductor. *NCP4624 [online]*. [cit. 2018-01-01]. Available from: http://www.onsemi.com/pub/Collateral/NCP4624-D.PDF

[9] Petenyi, S. *Voltage regulator circuits, systems and methods for having improved supply to voltage rejection (svr)*. US Patent US9742270, 2017.

Design and Analysis of an Interdigital Capacitor in Microstrip Line Technology for Low Noise X-band System Applications

Kai Fan

Microwave Electronics Lab, University of Kassel, Wilhelmshoeher Allee 73, 34121 Kassel, Germany

KaiFan@student.uni-kassel.de and micel@uni-kassel.de

Abstract. *This paper introduces the design of an interdigital capacitor (IDC) with even finger numbers using RO4003C substrate with the thickness of 0.508 mm working on a certain operating center frequency from 6 GHz to 15 GHz. The influence of several physical parameters such as width of fingers, length of the IDC, distance between adjacent fingers, space between finger end and interconnection part and the number of finger pairs are observed to conclude the final design of an IDC. With the feature of each parameter on the performance of IDC, a table and several empirical formulas are summarized to give recommendations about the calculation of the initial dimensions of IDC. The simulations are based on the model designed in ADS (Advanced Design System), which is a microwave simulation tool, using the MoM (Method of Moments). Using the linecalculator function in ADS, the width of the feedline with the impedance of 50 Ohm is calculated for respective operating frequency. The basic consistency between simulations and measurements, using de-embedding procedures verifies the feasibility of the table and empirical formulas.*

Keywords

Interdigital Capacitor, Advanced Design System, Method of Moments, Microstrip Line, Linecalc, De-embedding.

1. Introduction

In the design of RF integrated circuits, capacitors are one of the basic components. The most commonly used integrated capacitor in MICs (Microwave Integrated Circuits) is the Interdigital Capacitor (IDC). IDCs play an extremely important role, not only in microwave integrated circuits, but also in devices such as slow-wave devices, acoustic sensors, and converters [1].

For IDCs, various analytical design methods and ideas have been proposed [1-3]. The IDC was originally proposed by Gary D. Alley in 1970 [2]. His model is based on the lossless coupling line theory. But it does not take into account, the influence of discontinuities in the structure on the capacitance characters [2]. In addition to the above method, Zhu and Wu made some attempts to model the IDC from the perspective of admittance inverter [3]. However, the limitation of this method is that it does not take into account the high frequency loss. All calculations are based on the assumption that the network is a lossless network [3]. This assumption makes the analysis result still unable to accurately reflect the actual characteristics of IDC. The above mentioned methods are all researched in a relatively low frequency range and have their own drawbacks. Based on a full-wave analysis method (MoM) in Momentum from [4], this paper studies the characteristics of IDC in microstrip line technology on a higher frequency range, especially in X-band by creating the IDC model in the simulation software ADS (Advanced Design System). It will focus on the influence from physical parameters of IDC, such as the finger width to the S-parameters of IDC for different working frequencies. IDC, with a series of different size parameters is simulated and several empirical formulas are summarized, which are very useful for the initial design parameters of IDC.

2. Design of the IDC

The basic structure of simple and classic IDC with 6 fingers (3 finger pairs) is shown in Fig. 1. The two feed lines are connected to the interconnection part as shown in Fig. 1. The width of the fingers (w), the length of IDC (l), the gap between the fingers (g), the finger length (l_f), the width of the interconnection part w_t and the gap between the finger ends and interconnection part (g_e) are the physical parameters of IDC which influence the characteristics of IDC.

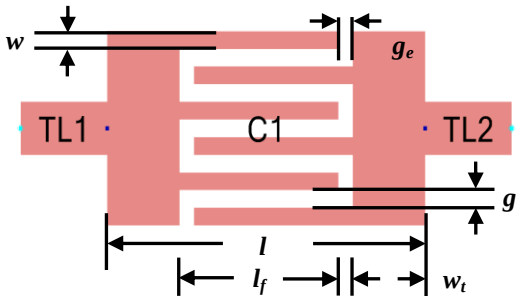


Fig. 1. Basic structure of an Interdigital Capacitor in ADS

Each finger width and finger length are equal, the width of the two interconnection parts is equal, the distance between two fingers and the gap between the finger end and interconnection part is equal. The two feed lines are equal in width, as the feedlines with different width [5] result in different impedances from feedlines. In this paper, the whole system is set to be a 50 Ohm system in order to match the impedance.

2.1 Simulation of the initial size of IDC

The simulations are done using the simulation tool ADS. All the physical parameters of IDC are analyzed in ADS and some conclusions are made from the analysis. Using the linecalc function in ADS, the width of the feedline with the impedance of 50 Ohm for a given center frequency is calculated.

IDC exhibits parasitic inductance caused by mutual inductance from the fingers, series resistance caused by the skin effect and proximity effect of the metal strip and dielectric loss and parasitic shunt capacitance to the ground [1].

After simulations for each parameter, the following conclusions are made: The width of the finger, space between fingers and the number of the fingers are wanted to be small in order to get better magnitude of S_{21}, where the gap between fingers and the number of fingers has significant influence on the magnitude of S_{21} in the passband. From the simulations, the change of the width has only small influence on the shift of the resonance frequency, while the change of the gap between fingers has a significant influence on the shift of the resonance frequency. It behaves like, the smaller the gap between fingers, the wider the bandwidth. At the same time, the overall width of the IDC is required to be greater than or equal to the width of the feed lines. Therefore, during the design, the minimization of the distance between the fingers and the number of the fingers must take precedence over the minimization of the finger width. Also, from the simulations, the space between finger end and interconnection part has almost no influence on the performance of IDC. In order to reduce the number of unknown parameters, it has been set equal to the gap between fingers.

From simulations, it has also been concluded: The shorter the length of the finger of IDC, the higher the resonance point in frequency. This means, that with the decrease of the length of the finger, the area of the cross-section of copper foil is also decreased. Therefore, the resonance frequency increases, as the capacitance is decreased. So initially, the determination of the working center frequency is required according to the application requirements. If the total width of IDC is larger than the width of the feedline, there is always a discontinuity as bend [6] at the contact surface between IDC and feedlines. Due to the loss from fringing field, this will affect S_{21}. The most important thing is to eliminate the discontinuities at the connection part between feedline and IDC, or at least to reduce the effect. In this work, the discontinuities at the contact surfaces between IDC and feedlines are eliminated by making the total width of IDC equal to the width of the feedline.

After eliminating the discontinuity, IDC has a better performance on the magnitude of S_{21} with less finger numbers. To satisfy the condition that the total width of IDC must be equal to the width of 50 Ohm feed line at a certain center frequency. Because the minimum size of physical parameters of IDC for the fabrication is 0.1 mm due to the precision and repeatability of the laser machine of Photolaser R from LPKF. Also, the width of feedline should be from 1.1166 mm to 1.1314 mm for frequencies from 6 GHz to 15 GHz from the calculation from linecalc, respectively. This means, the width of the feedline is greater than 1.11 mm but less than 1.13 mm. In order to let the total width be equal to the width of the feedline, the width of the feedline can be set 1.12 mm from 6 GHz to 12 GHz or 1.13 mm from 13 GHz to 15 GHz. Under the conditions above, the best option of finger numbers f_n is 4 under the condition that the gap between fingers is better as small as possible. There is still one physical parameter w_t, which is the width of interconnection part. But since the total width of IDC is equal to the width of feedline, the interconnection part can be accounted as the extension of the feedline, which has no influence on the performance of IDC. With the determination of width of feed line, the width of the fingers can be obtained by calculation from the width of the feedline.

After the analysis made from the simulations that are mentioned above, a table and several empirical formulas are made to give suggestions to get the initial sizes of IDC with the center frequency range from 6 GHz to 15 GHz as follows.

f_c/GHz	f_L/GHz	f_H/GHz	l_f/mm	p_b/GHz	C_1/pF	C_2/pF
6	5.22	6.77	8.08	1.55	0.104	0.95
7	6.08	7.88	6.9	1.8	0.089	0.82
8	6.94	8.99	6.01	2.05	0.078	0.7
9	7.82	10.13	5.3	2.31	0.069	0.59
10	8.7	11.28	4.73	2.58	0.062	0.52
11	9.6	12.44	4.26	2.84	0.056	0.48
12	10.49	13.58	3.87	3.09	0.051	0.41
13	11.35	14.7	3.55	3.35	0.047	0.38
14	12.21	15.8	3.28	3.59	0.043	0.33
15	13.12	16.96	3.03	3.84	0.04	0.3

Tab. 1. Recommendations of the initial size of IDC and the capacitances of IDC and an ideal capacitor

In Tab. 1, f_c is the center frequency, f_L is the lower cut-off frequency and f_H is upper cut-off frequency with magnitude of S_{21} equal to -3 dB, l_f is the length of the finger of the IDC, p_b is the bandwidth of the passband for each center frequency, C_1 in pF is the calculated capacitance of IDC from equation (1) [1] and C_2 in pF is the capacitance of an ideal capacitor, which has the almost

same performance within the passband from simulations. From the steps mentioned above, w, g, g_e, and f_n are fixed to a certain value with w equal to 0.205 mm from 6 GHz to 12 GHz or 0.2075 mm 13 GHz to 15 GHz, g, g_e equal to 0.1 mm *and* f_n equal to 4. These three parameters are not included in the table. Tab. 1 gives the recommended values for physical parameters of the IDC when the center frequency is in the range from 6 GHz to 15 GHz. ε_0 in equation (1) [1] is the vacuum dielectric constant, ε_{eff} is the effective dielectric constant, which is dependent on the operating frequency, and S is the area of the cross-section of the copper-foil. From equation (1) [1] the capacitance of IDC can be roughly calculated and its value is almost one tenth of capacitance from an ideal capacitor, which has the nearly same performance within the passband of IDC.

$$C_1 = (\varepsilon_0 * \varepsilon_{eff} * S)/g \quad (1)$$

According to Tab. 1, the formulas (2) and (3) to calculate the initial design parameters of IDC is formulated and the ratio of the bandwidth to center frequency is always approximately 26 % on the RO4003C with 0.508 mm. With the gap distance between fingers, the finger number and the width of the feedline, the width of the finger can be calculated. l_f is in meter and f_c is in Hz.

$$l_f = 3*10^{8} * f_c^{-1.073} \quad (2)$$

$$p_b = 0.259 f_c \quad (3)$$

In order to verify the feasibility of Tab. 1 and the formulas, three structures with center frequencies of 8 GHz, 11 GHz and 14 GHz are chosen to be fabricated and compared. Fig. 2 shows the actual structure of IDC from simulations for all the considered center frequencies. w_f is the width of the feedline.

The values of the design parameters for 8 GHz in Tab. 1 are taken to build this structure in ADS. The lengths of the feedlines on either side of the IDC structure are equal and for all the three structures, the total lengths of the two feedlines of the three IDCs are equal to maintain consistency while de-embedding.

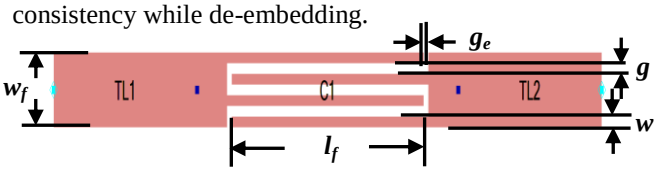


Fig. 2. The actual structure of IDC in ADS for 8 GHz

f_c/GHz	w/mm	l_f/mm	w_f/mm
8	0.205	4.13	1.12
11	0.205	2.38	1.12
14	0.2075	1.4	1.13

Tab. 2. The physical parameter values for the fabrication

Tab. 2 shows the actual values of the parameters from Tab. 1 for the fabrication of three IDCs. Other parameters were given before. The other two actual structures are like the form shown in Fig. 2. The values of the width of IDC in Tab. 2 are all calculated as mentioned before.

3. Fabrication and Measurements

The fabricated structure is shown in Fig. 3. From left to right are the IDCs with the center frequencies of 8 GHz, 11 GHz and 14 GHz, respectively, and a microstrip line twice the length of the feedline on each side of IDC. The space between two adjacent structures is 30 mm. The substrate used in the fabrication is RO4003C with the thickness of 0.508 mm because of its low loss and low production cost. The connector used in the measurement is 2.40 mm Jack (Female) End Launch Connector from Southwest Microwave.

The fabrication of microstrip feedline is for the de-embedding procedure. The de-embedding procedure is most needed here, because it is equivalent to removing the influence of the microstrip line on the performance of IDC. Fig. 4 shows the S_{21} from the measurement of the microstrip, which has twice the length of each feed line. It can be seen clearly that with increasing frequency, the loss becomes larger. So it is very necessary to eliminate the losses by de-embedding the IDC for all three center frequencies.

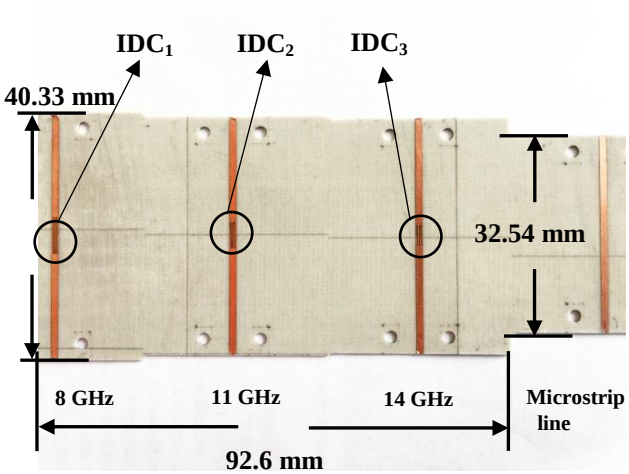


Fig. 3. The fabricated PCB with the structure of the IDC and the microstrip line

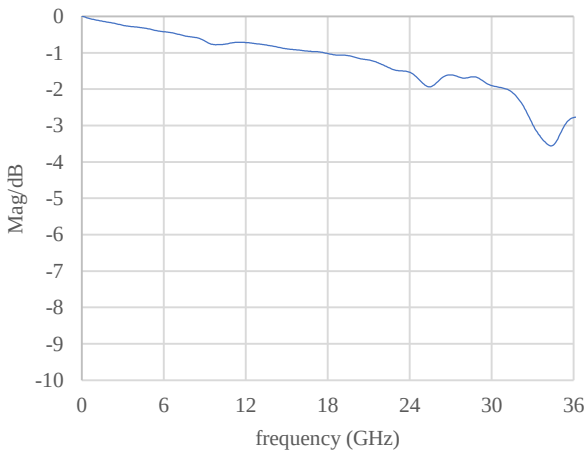


Fig. 4. $|S_{21}|$ from the measurement of the microstrip line

First, the IDC for 8 GHz center frequency is measured and the results of before and after de-embedding are shown

in Fig. 5. It shows that the simulation results are almost in agreement with the experimental results, the curves from the measurements show good bandpass feature and the bandwidth is very close to that from the Tab. 1. The resonance is at almost 16 GHz, of which the wavelength is twice as much as the quarter wavelength of 8 GHz. The reason for the occurrence of the resonance is, that each finger has inductance behavior and all the fingers are parallel connected to the capacitance between gap. From Tab. 1, with the help of the length of the finger, it can be concluded after calculations, that the length of the finger is equal to quarter wavelength. At 8 GHz, the magnitude of S_{21} is nearly 0 dB. At the same time, when the length of the finger is half wavelength of the resonance frequency, IDC is blocked. This means, when the length of the finger is equal to an odd multiple of a quarter wavelength, in the vicinity of its corresponding frequency, IDC shows good bandpass feature. On the other hand, when the length of the finger is equal to an even multiple of a quarter wavelength, IDC has blocking feature. As can be seen from Fig. 5, the second parallel resonance is near 34 GHz not 32 GHz. This is caused by the dispersion effect in [6]. The effective permittivity is frequency dependent. It increases as the frequency is increasing. When the frequency approaches positive infinity, it is equal to the relative dielectric constant [6]. Therefore, the effective permittivity is always smaller than the relative dielectric constant. The capacitance of IDC is indirectly influenced and it shifts the resonance frequency a little bit to high frequency. The effective permittivity is derived to be 3.44 in order to match the simulation to the measurement.

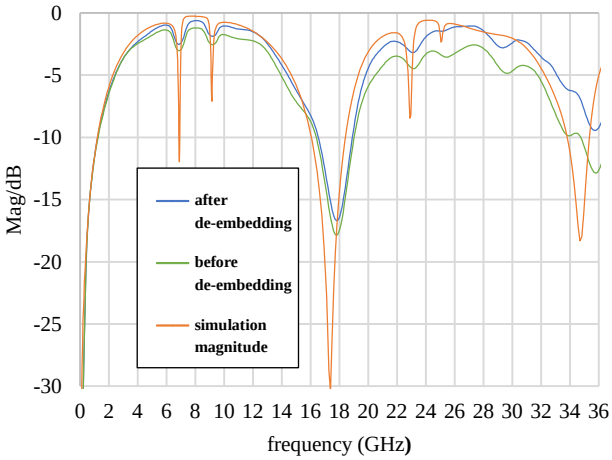


Fig. 5. Simulation and measurement results of $|S_{21}|$ for IDC1 (before and after de-embedding) with a center frequency of 8 GHz

As can be seen from Fig. 6, it shows similar behavior to the curve of Fig. 5. The effective permittivity is derived to be 3.47. The bandwidth is increased when compared to the previous measurement done for 8 GHz. The measurement shows that the structure shows good agreement to the simulation near the center frequency 11 GHz. Also, the resonance frequency is around 24 GHz, which is almost the double of the quarter wavelength of 11 GHz. The shift of the resonance frequency also shows the influence from the dispersion effect.

From Fig. 7, it is very obvious to see that the difference of the amplitude between the two measurements is almost the same compared to the other two situations. The effective permittivity is derived to be 3.49. It increases with the increase of frequency. The skin depth is inversely proportional to frequency. With the increase of frequency it decreases. This means, the area of the cross section, which the current flows through, decreases. Then the impedance of the microstrip increases. The difference is greater because of the larger conductor loss caused by skin effect like the previous two measurements, it shows good bandpass feature at 14 GHz and the passband width is increased through the comparison of three graphs as Tab. 1 and the formulas shows. The first resonance frequency is approximately 31 GHz not far from 28 GHz, which is twice as much as 14 GHz. This deviation is also derived from the dispersion effect. The previous conclusion helps to analyze the S parameters of IDC and verify roughly the correctness of the S-parameter from simulation and measurement.

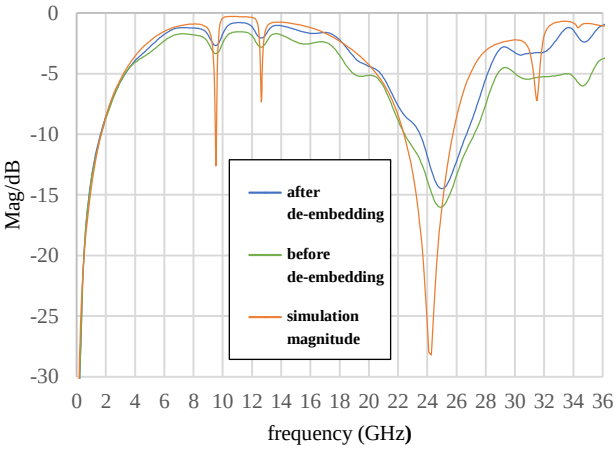


Fig. 6. Simulation and measurement results of $|S_{21}|$ for IDC2 (before and after de-embedding) with a center frequency of 11 GHz

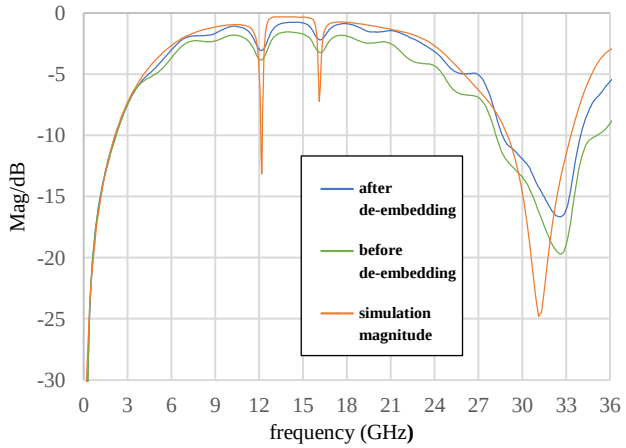


Fig. 7. Simulation and measurement results of $|S_{21}|$ for IDC3 (before and after de-embedding) with a center frequency of 14 GHz

Due to the change of the length of the finger, the area of the cross-section of finger is also changed. It influences the resonance frequencies of the IDC The length of the

finger is the key parameter to change position of the passband from IDC.

It can be seen from the comparison of the above three groups of data that measurement data is in good agreement, compared to the data from the simulation. It is also very obvious, that the differences of the magnitude between the simulation and the measurement are also significant although the de-embedding procedure is done. Radiation losses [6] are inevitable. With the increase of the working frequency, both the dielectric loss and conductor loss [6] due to skin effect are much higher.

4. Conclusions

With the help of the simulation, the influence of various physical parameters of IDC with the substrate RO4003C on the transmission characteristics of IDC is studied. After lots of simulations, it is concluded that with the decrease of width of finger and space between fingers, the performance of IDC is getting better and better. The gap end of IDC has little influence on the S_{21} of IDC and the length of the finger of the IDC can change the position of the passband. And as the number of the fingers of the IDC is smaller, the magnitude of S_{21} of IDC is much closer to 0 dB. After multiple MoM simulations of IDC, finally a table and some empirical formulas to calculate the initial size of each physical parameter of IDC are obtained.

By comparing the simulation diagram with the measurement diagram of three IDCs for three chosen frequencies, the results of measurements are basically consistent with the simulations. The feasibility of the table and empirical formulas has been confirmed. By comparing the capacitance of IDC and an ideal capacitor, it can be concluded, that the capacitance of IDC is one tenth of capacitance of the ideal capacitor, where this capacitor has the same performance in their passband. Also, the finger length changes the position of the resonance frequency and is always equal to a quarter wavelength of the operating frequency. The IDC is blocked, when the finger length is equal to an even multiple of quarter wavelength of the operating frequency. As mentioned earlier, it is important to eliminate discontinuities at the connection part between feed line and IDC or at least reduce the effect by taking the total width of IDC equal to the width of the feedline. In addition, two tapers [6] can be added at both sides of IDC to reduce the discontinuity effect, when IDC has more fingers. On one side, the taper width is matched to the total width of IDC and on the other side the width of taper equals to the width of 50 Ohm feedline. But this still needs to be verified. This points out the direction for future research work.

Acknowledgements

The research described in the paper was supervised by M.Sc. V. Kilaru, Microwave Electronics Lab, University of Kassel, Germany. I would like to thank Mr. V. Kilaru for his patience and meticulous guidance and Mr. C. Sandhagen and Prof. A. Bangert for his support during this work.

References

[1] Zhang Y., *Application of Artificial Neural Network in Analysis and Synthesis of Interdigital Capacitor*, 2005, Nan Jing in China: East South University.

[2] Alley G. D., *Interdigital Capacitors and their Application to lumped element microwave integrated circuits*, IEEE Trans. Microwave Theory and Tech., 1970, vol. MTT-l8, pp. 1028-1033.

[3] Zhu L., Wu K., *Accurate circuit model of Interdigital Capacitor and its application to design of new quasi-lumped miniaturized filters with suppression of harmonic resonance*, IEEE Trans. Microwave Theory and Tech, 1998 vol. 8, no. 3, pp. 130-132.

[4] Chen Y., *Application of Method of Moments in RF circuit design*, 2006, Wu Han in China: Huazhong University of Science and Technology.

[5] Fengliu X., *Study on Mechanism and Application of Microwave Interdigital Structure*, 2005, Su Zhou in China : Su Da Yuan

[6] Kompa G., *Practical Microstrip Design and Applications*, 2005, Boston: Artech House.

[7] Halliday D., Resnick R., *Fundamental of Physics*,-6th ed. New York, John Wiley &Sons, Inc., 2001

About The Author...

Kai Fan was born in 1990. He is doing his Master's degree in Electronical Communication Engineering from University of Kassel. His major focus is microwave electronics. His research interests are antenna design and microwave integrated circuits design.

Audio Quality Assessment of Digital Radio Broadcasting Systems

Adam BARTYZAL

Dept. of Electromagnetic Field, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic
bartyada@fel.cvut.cz

Abstract. *Due to the amount of content in digital radio systems it is not possible to determine the sound quality of the source coding by subjective tests. In the last two decades, systems of objective evaluation using auditory models have been developed. With the development of more modern codecs utilizing psychoacoustic knowledge, machine methods seem to provide inaccurate results, i.e. they give worse quality scores than the human listener, especially at low bit rates. The paper focuses on two methods of evaluation: PEAQ and ViSQOL, their evaluation of different coding methods depending on the bitrate and tries to determine whether they are undervaluing quality and to what extent.*

Keywords

audio quality assessment, codecs, PEAQ, PEMO-Q, ViSQOL, MUSHRA

1. Introduction

Frequency spectrum is a very expensive matter in these days, and with the development of telephone data services, the bandwidth designated for radio broadcasting is decreasing. In contrast, the amount of content offered is growing. The solution to combine these two trends is digital radio DRM or DAB. It is necessary to find the optimal level of compression of the source signal so that the listener does not realize there is a difference from the uncompressed original and at the same time the content provider uses the spectrum as effectively as possible. To determine such a threshold, subjective tests might be used, however, their time-consumption (therefore financial cost) is high and because of that automation and computer algorithms are being introduced to provide an objective assessment.

2. Objective evaluation methods

In listening tests of signals encoded by different codecs, one of the two following methods is usually used being according to ITU-R recommendations [1]. ABX (Double Blind Test) and MUSHRA (MUltiple Stimuli with Hidden Reference and Anchor). The outputs of both testing methods can be mapped to a scale defined in [1] called SDG (Subjective difference grade). Objective algorithms provide their results on the ODG (Objective difference grade) scale that corresponds to the SDG as shown in table 1.

Impairment	Grade	SDG	ODG
Imperceptible	5	0	0
Perceptible, but not annoying	4	-1	-1
Slightly annoying	3	-2	-2
Annoying	2	-3	-3
Very annoying	1	-4	-4

Tab. 1: Quality grades defined by [1]

2.1 PEAQ

Perceptual Evaluation of Audio Quality defined by Recommendation ITU-R BS.1387 [2] exists in two different versions of the algorithm. Basic and advanced. There are several differences between them. The base version does not use a bank of filters to model the ear and the size of the neural network matrix used to calculate ODG.

PEAQ requires a full reference, that is, the original signal (not distorted by compression), which is compared with the test signal. Both signals pass through the psychoacoustic model, whose output are MOVs (Model Output Variables). These are then used by the artificial neural network to calculate the results as shown on Fig. 1. Detailed description of the algorithm can be found in [10].

Basic version can be obtained for free, implemented in Matlab (available at [3]). Advanced version is implemented in software called Opera [11] by a company Opticom who participated in development of the recommendation. Advanced version is being used in this article.

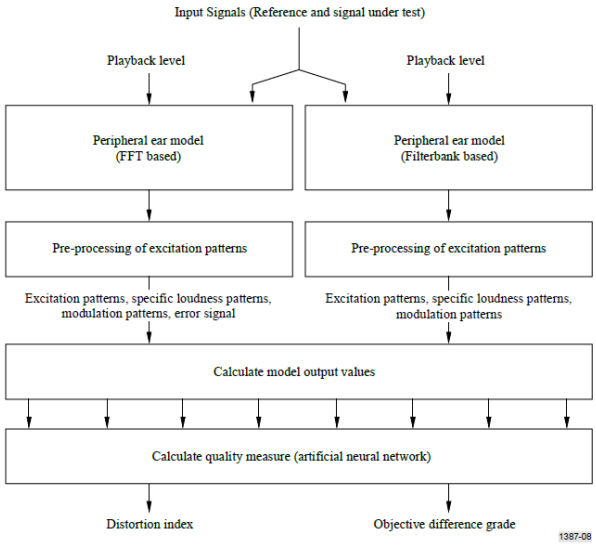


Fig. 1. Data flow in PEAQ Algorithm [10]

2.2 ViSQOL

Unlike the previous method, an algorithm called ViSQOL (Virtual Speech Quality Objective Listener) does not try to determine the amount of distortion, noise, etc. Originally the algorithm was used for assessment of speech quality in telecommunications. Later some adjustments were presented in [1] in order to evaluate quailty of music signals. On Figure 2 is shown how the algorithm operates.

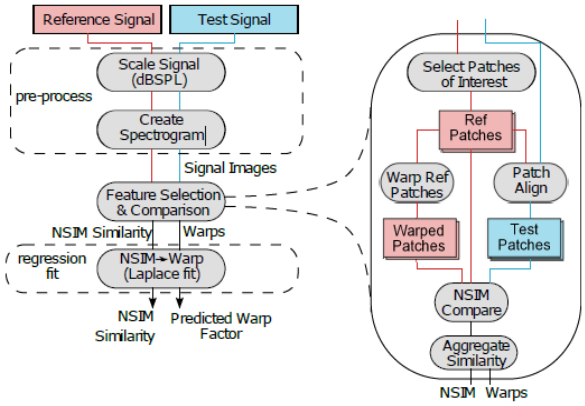


Fig. 2: Block diagram of ViSQOL Algorithm [1]

According to [4] similary to PEAQ, ViSQOL compares the distorted signal with the original, but it does it differently. In the pre-processing phase both signals are being scaled to same level. Second part of pre-processing is calculation of spectrograms using Short Time Fourier Transform STFT. Thirty logarithmically distributed filters from 250 Hz to 8 kHz simulating basilar membrane are used when assessment of speech is chosen. When evaluating musical samples, the whole hearing band is covered.

Second part of the algorithm called *Feature Selection and Comparison* takes care of selecting so-called "Patches of Interests". Patch is a part of spectrogram with length of thirty frames. In speech evaluation three patches in bands 250 Hz, 450 Hz and 750 Hz with the highest intensity are being chosen for Mean Opinion Score computation. ViSQOLAudio overcomes this selection and uses all the patches.

Lastly spectrograms are converted to neurograms (time-neural firing activity dependencies). They are stored in form of images and then compared using NSIM (Neurogram Similarity Index Measure).

2.3 Reference files

Because testing should reflect the actual radio broadcast, samples representing different musical genres and spoken words were selected.

At the same time, these samples were adapted for subjective testing, i.e. they were rimmed to ten seconds in duration and provided with one second fade in and fade out. List of audio samples is shown in Tab. 2.

#	Filename	Description	Type
1	capriccio.wav	Capriccio Italien Op. 45	music
2	cimrman.wav	Part of Czech play "Opeřený Had"	speech
3	dubstep.wav	Electronic music	music
4	holmes.wav	Reading from a book: Sherlock Holmes	speech
5	pennylane.wav	The Beatles: Penny Lane	music
6	rickroll.wav	Rick Astley: Never Gonna Give You Up	music

Tab. 2 Set of audio files

2.4 Used Audio Codecs and bitrates

Coding of previously mentioned audio samples was prepared to standard used in DAB/DAB+ [5]. According to (mpeg) MPEG 1 Layer II can be coded into fourteen possible bitrates from 8 kbps to 384 kbps. Encoder TwoLame [6] was used set to stereo output.

Second coding method used was Advanced Audio Coding in three different profiles. First LC-AAC (Low complexity) is based on conversion to frequency domain using Discrete Cosine Transform similarly to JPEG. HE-AAC v1 adds to LC a process called Spectral Band Replication. Since last one or two octaves of audible spectrum don't carry an important information but when they are simply cut off by low pass filter listeners describe sound as "dull". SBR replicates lower part of spectrum to simulate higher harmonics. The third used profile was HE-AAC v2 which adds the usage of Parametric Stereo to previous profiles, combining stereo channels into one mono channel. The Spatial information is coded into three parameters Inter-channel Intensity Difference (IID), Inter-channel Cross-Correlation (ICC) and Inter-channel Phase Difference (IPD). Samples were encoded with FFMpeg [7] in following bitrates. For LC profile from 8 kbps to 256 kbps with step 4 kbps. For v2 profile the same only the upper limit is limited to 128 kbps and for v2 profile to 64 kbps.

3. Results

Fig.1 shows a comparison of two selected methods. In the first line, the ODG averages of the speech samples are plotted and the second is the average of the music samples. The first column shows the PEAQ rating. For better comparison of individual codecs, the dependencies are plotted on a single graph. The second column shows the results of the ViSQOL algorithm. Its standard output is MOS-LQO, which corresponds to ratings on a scale of one to five. In comparison, however, MOS-LQO is remapped to ODG. The last column shows the difference between the two methods. It is obvious that the biggest differences are at low bitrates, which is already suggested by articles [8] and [9]. ViSQOL gives higher quality to more modern codecs.

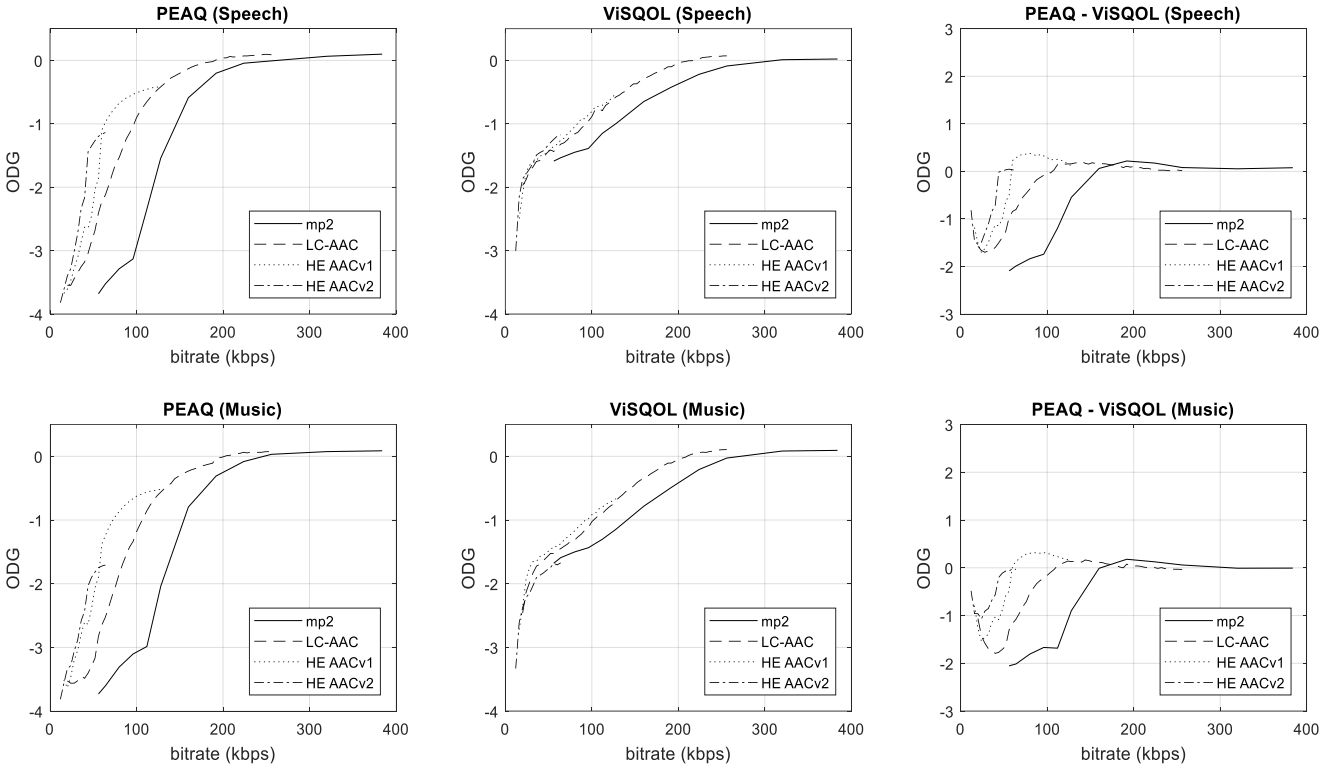


Fig. 2: Left: Average ODG evaluated with PEAQ, Center average ODG evaluated with ViSQOL, Right: Difference between them

4. Conclusion

As shown in the third column, the difference in evaluation of the two presented methods reaches up to two degrees on the ODG scale. Which one of the methods represents reality more precisely must be judged by subjective methods. The MUSHRA [2] was selected to verify this credibility. An interface for the test was prepared in Matlab (shown on Fig. 2), and preliminary results suggest that ViSQOL is closer to reality. However, it is necessary to perform the tests on a larger group of subjects, which the author will devote to in his diploma thesis.

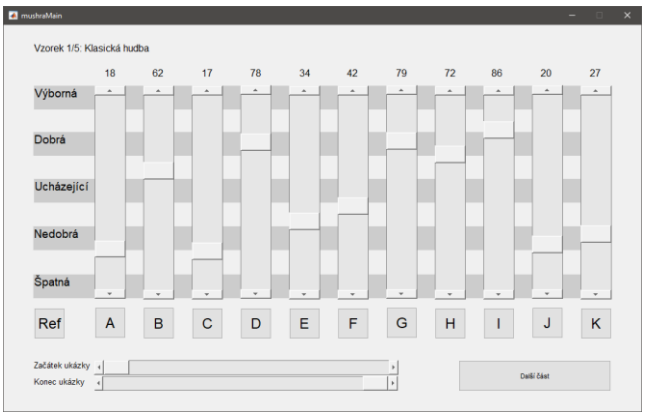


Fig. 3: GUI of MUSHRA test

Acknowledgements

Research described in the paper was supervised by Ing. K. Ulovec Ph.D., FEE CTU in Prague and supported by the Grant Agency of the Czech Technical University in Prague, under grant No. SGS17/190/OHK3/3T/13.

References

[1] ITU-R, „Methods for the subjective assessment of small impairments in audio systems“. Recommendation ITU-R BS.1116-1, 1997.

[2] ITU-R, „Method for the subjective assessment of intermediate quality level of coding systems“. Recommendation ITU-R BS.1534-1, 2003.

[3] P. KABAL, „GitHub,“ 2002. [Online]. Available: https://github.com/NikolajAndersson/PEAQ.

[4] J. MCNALLY, Towards improving ViSQOL, Dublin: Dublin Institute of Technology, 2017.

[5] E. T. S. INSTITUE, „Digital Audio Broadcasting (DAB)“. Recommendation ETSI TS 102 563 V2.1.1, 2017.

[6] M. CHENG, „TwoLame,“ [Online]. Available: http://www.twolame.org/.

[7] FFMPEG PROJECT, „FFMpeg Documentation,“ 2019. [Online]. Available: https://ffmpeg.org/documentation.html.

[8] F. RUND a K. ULOVEC, „Comparison of Two Objective Methods of Quality Assessment for Digital Audio Broadcasting,“ *Radioelektronika,* 2018.

[9] P. POČTA a J. G. BEERENDS, „Subjective and Objective Assessment of Perceived Audio Quality of Current Digital Audio Broadcasting Systems and Web-Casting Applications,“ *IEEE TRANSACTIONS ON BROADCASTING,* sv. 16, č. 3, 2016.

[10] International Telecommunication Union. Rec. ITU-R BS.1387-1: Method for objective measurements of perceived audio quality. ITU-R, 2001, 100 p

[11] Opticom, „OPERA™,“ Opticom, 2019. [Online]. Available: https://www.opticom.de/products/opera.php.

About Author...

Adam BARTYZAL was born in Jindřichův Hradec, Czech Republic in 1994. In 2016 he obtained his bachelor’s degree in department of Electromagnetic field on Faulty of Electrotechnics, Czech Technical University in Prague and currently is continuing in master’s degree phase in the same department.

All-Inkjet Printed Flexible Capacitors Utilizing SU-8 and PVP Dielectric Inks

POVOLNÝ Vojtěch[1], *PUCI Florian*[2]

[1] Dept. of Telecommunication, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

[2] Dept. of Microelectronics, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic
povolvoj@fel.cvut.cz, puciflor@fel.cvut.cz

Abstract. *This paper presents fabrication of printed interdigital capacitors on a flexible and lightweight polyester (PET) substrate. Capacitors were fabricated using inkjet printing technology and conventional low temperature (100-150°C) curing method. The surface of printed capacitors were covered with dielectric material, which lend capacitor a specific value of permittivity, but also serves as protective coating. Cross-linked poly-4-vinylphenol (cPVP) and SU-8 were applied as dielectric materials and experimental results are compared.*

Keywords

Inkjet, printed electronics, passive electronic components.

1. Introduction

The inkjet technology is known for the ability of a high precision used when depositing a liquid in a well-defined and pre-designed digital pattern. This makes it an essential and more popular technology for fabricating electronic devices and many applications. It is a computer printing technology and the main advantages are the low cost, minimal waste generation and efficient handling of expensive materials. Therefore, it is suitable for development of new technologies or patterns. It´s also relatively fast technology of fabrication of samples. And because there is no wasting of material it is more financial efficient.

In this article, inkjet printing is primarily used for the printing of capacitors. The electrical characteristic of the capacitor is strongly dependent on the morphology of the printed layers, the material used for printing and every printing parameter.

Referring to the manner how the electrodes are arranged in a capacitor, there exist two basic types – co-planar and parallel plates. Each of these types has a different and specific process of fabrication. A well-known and common type capacitor, consists of two parallel plates from conductive material, among which is placed the dielectric material. The capacity is then depending on the area of the plates, on gap between electrodes and on the permittivity of the dielectric material being used.

However, the production of this type by inkjet printing is very difficult, because it requires a specific dielectric material which would be used and has the ability to print the top electrode without problems (pinholes not formed in the dielectric film which would short circuits). Part of the production process requires that each material must be sintered in an oven and this requires precise positioning of the samples in order to fit next upper printed layers to correct place. Finally, it makes the production time expensive. [1-4]

The second type is a co-planar capacitor. That means that both electrodes are lying next to each other and are covered with dielectric material. To increase the value of capacity, co-planar capacitors have fingers which increase the area of electrodes and thus the capacity. Fabrication of this type is not so difficult, because there are only two layers of material which needs to be printed and therefore it is less time consuming. However, to achieve capacitance values of 1pF and more, then it is necessary to increase the printed are of the capacitor.

This type has another more utilization. Covering the pattern with sensitive material it could serve as a sensor of gases or humidity. A planar interdigital sensor functions based on the rule of two parallel plate capacitors, where electrodes open up with the electric field lines surrounding the material under test. The parallel in-plane electrodes create a digit-like or finger-like periodic pattern and that's why the ''interdigital'' term is used. This electric field generated by the sensor has the ability to change the impedance. A response to the interaction between the layers provides the ability to measure the change in capacitance or impedance. The capacitive reactance is a function of system properties and with the sensor behaving similarly to a capacitor, these properties can be evaluated by measuring the capacitance. The output would be a high signal-to-noise ratio impacted by the coplanar positioning of the electrodes and stronger signal can be achieved by repeating the pattern.

Combination of printed capacitor and resistor can be created all-inkjet printed RC circuit as can be seen on Fig. 3. There also can be seen Amphenol FCI Clincher™ Flex connector, which makes measurement much easier and does not cause damage of structure.

Substrate for printing can be selected a lot of materials. As examples, let's mention Polyethylene terephthalate (PET), Polyethylene naphthalene (PEN), Polyimide (PI), or Photo paper. What material will be used depends on thermal condition of ink, substrate and purpose of usage of final product. To protect printed layers/patterns can be used transparent ink or for better protection is good to use thin lamination foil well known from lamination of paper documents.

2. Fabrication method

Fig. 1. shows the schematic representation of the fabricated capacitor. Whole pattern is printed on PET foil which has one side pretreated for printing. On the top of the meander is printed capacitor is covered from a dielectric material. Every capacitor have height of 20 mm. Dimensions can be seen in Tab. 1.

Substrate preparation:

As mentioned earlier, the substrate used for printing was pretreated from the production, so therefore it has ideal conditions for inkjet printing.

Printing:

The inks (silver nano-dispersion from Sigma-Aldrich PN-736465 and XP PriElex SU-8 polymer material from Microchem) were printed with the inkjet printer Fujifilm Dimatix DMP-2831. The platen temperature was set to 40 °C for silver ink and to 30 °C for SU-8 ink and the cartridge to 35 or 31 °C respectively. The drop spacing was set to 25 µm (1016 dpi), jetting frequency 2 kHz, jetting speed 11 m/s, as can be seen in Fig. 2. Curing was performed in laboratory oven at 140 °C for one hour. The printing result is shown in Fig. 3.

Further, I would like to mention that c-PVP dielectric material was prepared at laboratory in our university. The preparation requires the following steps.

- 10 ml of propylene glycol monomethyl acetate (PGMEA)
- Poly(melamine-co-formaldehyde) methylated (PMFM), 84 wt% in 1-butanol as a crosslinking agent
- The weight ratio of PVP to PMFM - 5:1
- Before printing, the PVP-based dielectric solution dilute in PGMEA (volume ratio 1:1) and filter with a 0.2 µ m syringe filter to remove residual agglomerations

3. Experimental results

Measurement of the capacity was performed for all the printed samples. As can be seen in Tab. 2., the measured values of capacity are in pF. Firstly, the samples with no added dielectric layer were measured. It was proven that doubling the printed area of the capacitor, it increases the value of capacity by almost twice. However, the absence of the dielectric material would question the stability of the measured values. The resulting capacity is affected by air and water vapor. Next measurement was carried out with dielectric material SU-8. In this case the capacity values increased by 21 % against the values where air was used as dielectric material. Still was maintained that the capacity raised by twice of its value while printing twice larger area of the capacitor. The other dielectric material used was c-PVP. In this case the capacity increased by about 0.8 % against the samples using SU-8. And still the capacity increased by using larger printing area of the capacitor. All values can be seen in Tab. 2 and are depict in Fig 5 and Fig 6. Next measurement was applied to sample with length and width of fingers set to 10 mm and the gap between the capacitor fingers was 500 µm. It has been investigated that there is an influence of adding layers of dielectric material. In Tab. 3 are shown the capacity values for samples with c-PVP layers from 0 up to 3 layers. It can be seen that the greatest impact is just by adding dielectric material. Adding more layers of the dielectric material does not influence much the values of the capacity.

The last measurement was applied on printed RC circuit, which is depicted on Fig. 4. Parameters of component were set to 2.3 kΩ for resistor and 20 pF for capacitor, which correspond to 3.46 MHz limit frequency. Experimentally was found that resistor has value of 2.3 kΩ and capacitor 21 pF. I received limit frequency 3.5 MHz what fit our assumption.

Width of IDC fingers	Space between IDC fingers	Length of IDC fingers
500 µm	500 µm	5 mm
250µm	250µm	10 mm
		20 mm

Tab. 1. Dimensions of capacitors, which were printed in combination

	_S043	_S044	_S045
Air	3,75 pF	6,38 pF	11,82 pF
SU-8	4,58 pF	7,71 pF	14,49 pF
c-PVP	4,92 pF	8,38 pF	15,84 pF

Tab. 2. Values of capacity of a sample which has the same dimensions but different dielectric material

0 layers	1 layer	2 layers	3 layers
2,65 pF	2,83 pF	2,92 pF	2,98 pF

Tab. 3. Values of capacity depending on different count of dielectric layers

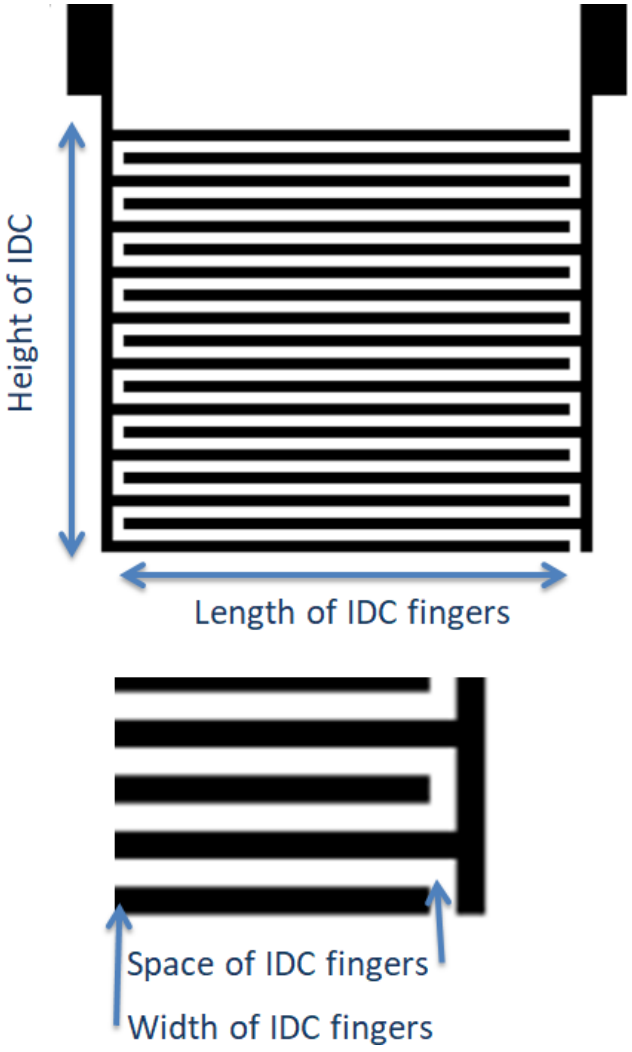


Fig 1 Dimensions of printed capacitors

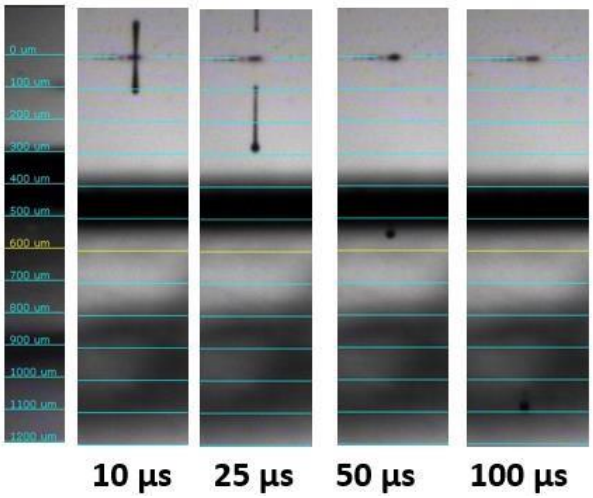


Fig. 2. Drop watcher window

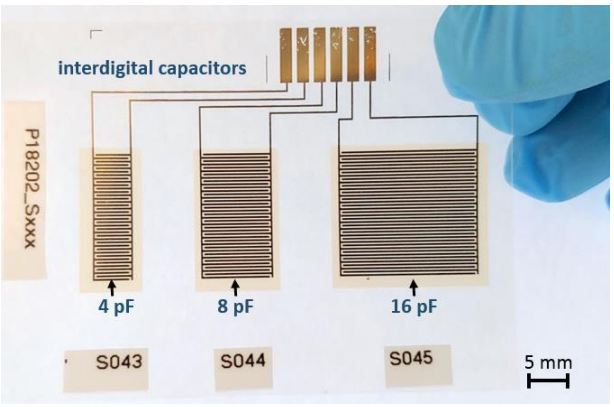


Fig. 3. Three types of capacitors have capacities of 4, 8 and 16 pF depending on the capacitor area dimensions. These three examples are covered with SU-8 dielectric material.

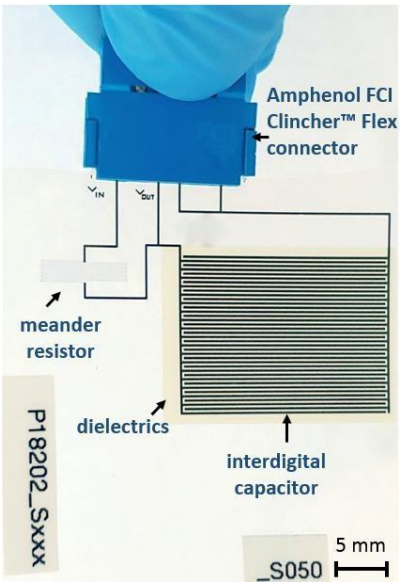


Fig. 4. All inkjet printed RC circuit with connector

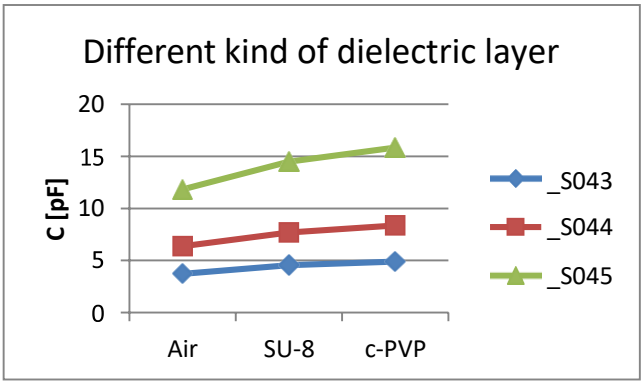


Fig. 5. Influence of dielectric material on capacity

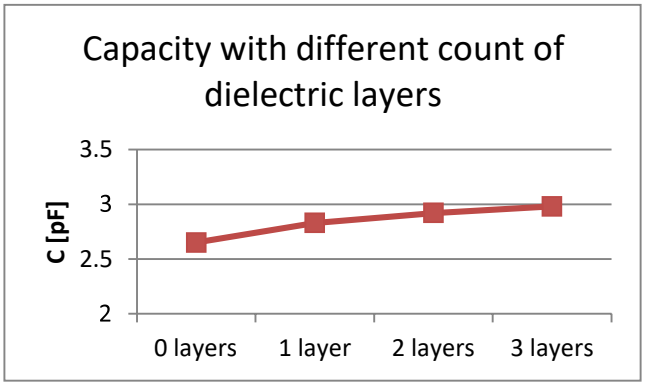


Fig. 6. Influence of number of dielectric layers

Conclusion

We established a printing process for fabrication of printed capacitors with dielectric cover material. Samples were subsequently characterized in terms of capacity, stability and repeatability. Then were studied influence of count of dielectric layer. There were obtained that 2 layers of dielectric material should be ideal. Further we printed RC circuit, for which was firstly counted limit frequency. By measuring was verified that we can print RC circuit with our desired properties. The results are still preliminary and further research will be continued.

Acknowledgements

This work is part of the CTU SGS grant no.SGS17/188/OHK3/3T/13 and was supervised by Ing. Alexandr Laposa and Prof. Miroslav Husák (FEE CTU in Prague). The authors wish to thank Assoc.prof. Jan Voves (FEE CTU in Prague) for many helpful advices.

References

[1] Y. Li, R. Torah, S. Beeby, and J. Tudor, "An all-inkjet printed flexible capacitor on a textile using a new poly(4-vinylphenol) dielectric ink for wearable applications," in *Proceedings of IEEE Sensors*, 2012.

[2] H. W. Tan, T. Tran, and C. K. Chua, "A review of printed passive electronic components through fully additive manufacturing methods," *Virtual Phys. Prototyp.*, vol. 11, no. 4, pp. 271–288, 2016.

[3] B. S. Cook, J. R. Cooper, and M. M. Tentzeris, "Multi-layer RF capacitors on flexible substrates utilizing inkjet printed dielectric polymers," *IEEE Microw. Wirel. Components Lett.*, vol. 23, no. 7, pp. 353–355, 2013.

[4] X. Hu and W. Yang, "Planar capacitive sensors - Designs and applications," *Sens. Rev.*, vol. 30, no. 1, pp. 24–39, 2010.

About Authors...

Vojtěch POVOLNÝ was born in May 1992 in Prague, Czech Republic. In February 2017 he received his Bachelor degree in Electrical Engineering from CTU in Prague (Prague, Czech Republic). He continued his master studies at CTU in Prague focusing on all inkjet printed capacitors on flexible substrates.

Florian PUCI was born on 10 July 1989 in Tirana, Albania. In the years 2004-2008 studied the "Sami Frasheri" high school in Tirana. In 2008 moved to Czech Republic to follow the preparation course for Czech language, which is part of the Charles University in Prague. In 2012, Florian PUCI graduated with a Bachelor Degree from the Czech Technical University in Prague in the study program Communication, Multimedia and Electronics, in the branch of Applied Electronics. After successfully completing the Master Study in the CTU of Prague in Electronics, in 2014 received the title Ing. Currently Florian is a Ph.D. student of the Department of Microelectronics in the Czech Technical University in Prague.

Radiation Tolerant 8-bit Analog to Digital Converter with Successive approximation

Pavel Vancura[1,2]

[1]Dept. of Microelectronics, Faculty of Electrical Engineering, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic
[2]Dept. of Physics, Faculty of Nuclear Sciences and Physical Engineering, Brehova 7, Praha 1

vancupa2@fel.cvut.cz

Abstract. *This paper brings implementation of 8-bit asynchronous analog to digital converter with successive approximation (SAR ADC) in 180 nm CMOS SoI technology. Radiation tests have proven that used technology is radiation tolerant up to 1 kGy. Layout in the used technology occupies 319 × 115 μm^2. The proposed SAR ADC consumes 295 μW from 1.8 V power supply at 4 MHz sampling frequency. Achieved ENOB is 7.81 bit and calculated figure of merit is 163 fJ/conversion-step.*

Keywords

radiation tolerant, SAR ADC, low power, asynchronous, fully-differential

1. Introduction

Modern electronics for special applications such as space applications, X-ray monolithic detectors, avionics, CERN experiments, etc. needs to be working in a radiation environment. Electronics in integrated circuits can be hardened by design, technology or layout techniques. This paper brings 8-bit SAR ADC in 180 nm SoI technology working in radiation environment up to 1 kGy. This maximum radiation value limits used technology [1]. The proposed design brings innovation of Harpe at al.[2] with improvements for radiation environment. The first improvement is the used SoI CMOS technology which is more radiation tolerant than classic bulk technologies [1]. The second improvement is using layout matched structures in combination with differential design which helps to eliminate single event effects (SEE). The third improvement is own customization of metal-oxide-metal (MoM) capacitor used in capacitor DAC of the SAR ADC. The figure of merit of the proposed design is 163 fJ/conversion step which is competitive value in comparison with existing published results, for example with references [3], [4], [5].

2. 8-bit SAR ADC circuit description

The 8-bit SAR ADC circuit design is described in detail by Harpe et. al [2]. In this section, only brief circuit description is provided and modifications are emphasized. A block diagram is shown in Fig. 1.

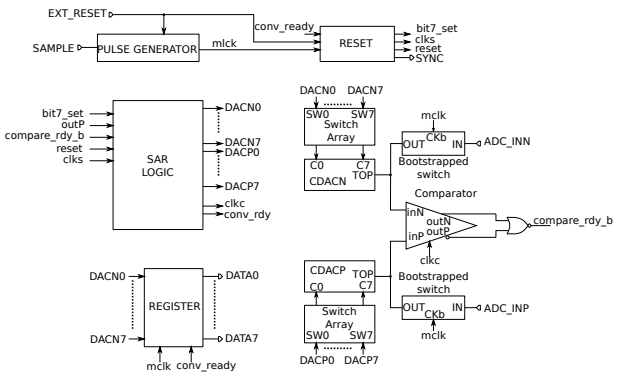


Fig. 1: 8-bit SAR ADC block diagram

The proposed design contains in comparison with [2] a pulse generator circuit which generates sampling pulse with adjustable length at rising edge of the SAMPLE signal. A bootstrapped switches [6] are used instead of transfer gates to improve the linearity of sampling signal into top plates of capacitor arrays. The output register latches data when conversion is finished. Data are latched until the next SAMPLE signal is received. Both of capacitor DAC's uses customized MoM capacitor shown in Fig. 2. Each of unit capacitor has dimensions 4.24x4.24 μm^2 with capacitance 4.5 fF. The proposed unit capacitor acts as a shielding box which improved linearity. A common centroid layout in the comparator is used. The common centroid layout in combination with fully differential ADC design mitigates SEE [7].

3. Results

Differential and integral non-linearities (DNL, INL) of the SAR ADC obtained from simulations are shown in Fig. 3 and Fig. 4. These nonlinearities have been extracted from

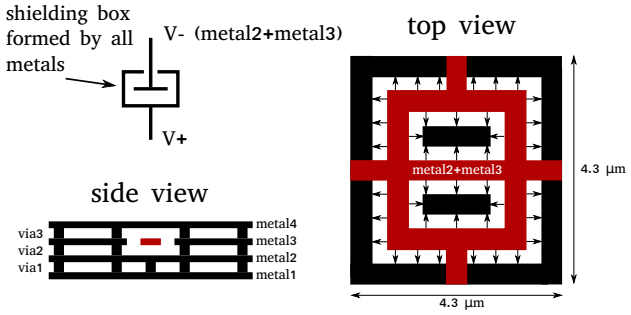


Fig. 2: The proposed MoM capacitor

transfer function by gradually increased step of 1 mV (0.25 LSB) at the ADC input. The worse case DNL is held within 1 LSB and -0.5 LSB and for INL between 1 LSB and -2 LSB.

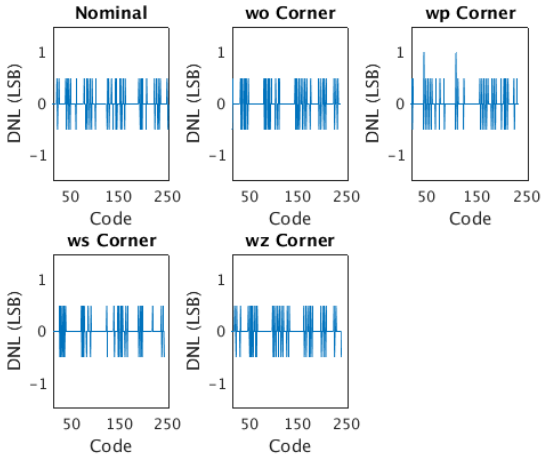


Fig. 3: Simulated DNL

The implemented prototype in 180 nm SoI CMOS technology is part of a monolithic pixel detector. For this reason, is impossible to measure SAR ADC directly and therefore only simulation results are provided. Measurements of the implemented prototype have revealed a problem with limited range of AD converter. The problem is caused by incorrect layout of PMOS voltage divider which sets reference of the fully differential amplifier which drives SAR ADC input. This failure resulting in the limited output voltage range of analog circuits driving the SAR ADC. The problem was understood and will be eliminated in future circuit re-design. However, the proposed SAR ADC seems to be working correctly.

Acknowledgements

This work is part of project Centre of Advanced Applied Sciences co-financed by the European Union with the number: CZ.02.1.01/0.0/0.0/16_019/0000778 and SGS grant with the number: SGS17/188/OHK3/3T/13.

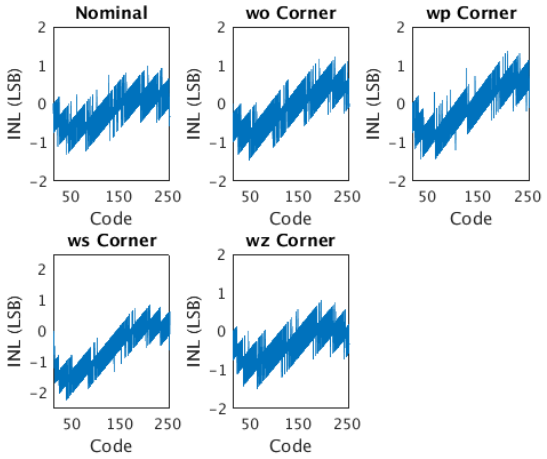


Fig. 4: Simulated INL

References

[1] Mariovsk, M.; Benka, T.; Havrnek, M.; Hejtmnek, M.; Janoka, Z.; Kafka, V.; Mariovsk, M.; Neue, G. et al. A comparative study of the TID radiation effects on ASICs manufactured in 180 nm commercial technologies Journal of Instrumentation. 2018, 13 ISSN 1748-0221.

[2] HARPE, Pieter JA, et al. A 26μ W 8 bit 10 MS/s Asynchronous SAR ADC for Low Energy Radios. IEEE Journal of Solid-State Circuits, 2011, 46.7: 1585-1595.

[3] BAGHBANMANESH, MohammadReza; MALOBERTI, Franco; GATTI, Umberto. A 10-Bit Radiation-Hardened by Design (RHBD) SAR ADC for Space Applications. In: 2017 New Generation of CAS (NGCAS). IEEE, 2017. p. 53-56.

[4] XU, Hongda, et al. A 78.5-dB SNDR Radiation-and Metastability-Tolerant Two-Step Split SAR ADC Operating Up to 75 MS/s With 24.9-mW Power Consumption in 65-nm CMOS. IEEE Journal of Solid-State Circuits, 2019, 54.2: 441-451.

[5] KUPPAMBATTI, Jayanth, et al. A radiation-hard dual-channel 12-bit 40 MS/s ADC prototype for the ATLAS liquid argon calorimeter readout electronics upgrade at the CERN LHC. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2017, 855: 38-46.

[6] RAZAVI, Behzad. The bootstrapped switch [a circuit for all seasons]. IEEE Solid-State Circuits Magazine, 2015, 7.3: 12-15.

[7] ARMSTRONG, S. E., et al. Demonstration of a differential layout solution for improved ASET tolerance in CMOS A/MS circuits. IEEE Transactions on Nuclear Science, 2010, 57.6: 3615-3619.

About Authors...

Pavel Vancura was born in 1981. Master's degree in the field of Microelectronics completed in 2017 at the Czech Technical university in Prague (CTU), faculty of electrical engineering (FEE). Currently he is a PhD student at CTU FEE and he is also working at the Faculty of Nuclear Sciences and Physical Engineering Czech Technical University in Prague, department of Physics, as a researcher involved in development of semiconductor pixel detectors.

Characterization of LVDS drivers for MAPS detector designed in 180 nm SOI CMOS technology

Tomas BENKA[1, 2]

[1] Dept. of Microelectronics, Faculty of Electrical Engineering, Czech Technical University, Technicka 2, 166 27 Praha, Czech Republic
[2] Dept. of Physics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Brehova 7, 115 19 Praha, Czech Republic

benkatom@fel.cvut.cz

Abstract. *The Low-Voltage Differential Signaling (LVDS) drivers for the X-ray imaging pixel detector X-CHIP-03 have been developed using a 180 nm deep submicron Silicon On Insulator (SOI) CMOS commercial technology. The X-CHIP-03 is a Monolithic Active Pixel Sensor (MAPS) and its architecture is based on the first prototype X-CHIP-02 with hit counting. Target application of the sensor is radiation imaging. The deep submicron SOI CMOS technology has long been used in many special application, such as radiation-hardened or high-voltage integrated circuits. It is only in recent years that SOI has emerged as a serious contender for low-power high-performance applications. The main feature of the SOI pixel detectors is the sensor location. A sensor element is implemented on handling wafer. The epitaxial layer which contains the front-end electronics is separated from the sensitive part of structure Burried-Oxide (BOX).*

Keywords

X-ray detectors, SOI CMOS technology, VLSI electronics, radiation damage to electronic components, LVDS drivers

1. Introduction

The applicability of the X-CHIP-02 has been found for many commercial applications, for example, the imaging detectors in the medicine industry, detectors for dosimetry and spectroscopy, tracking detectors, etc. However, all of them require to improve the speed of data acquisition from the detector. The communication speed and its radiation resistance is one of the most critical parameter of the detector. The X-CHIP-02 communication consists of long shift register daisy chaining all pixels. This architecture limits the maximum data rate below 50 kbps. The LVDS transceiver and receiver allow to increase the speed of communication to 500 Mbps with respect to low power consumption (up to 6 mW) and high radiation resistance. The motivation for development the LVDS IP block is the utilization of the detector in the commercial application. The LVDS receiver and transmitter are included in the X-CHIP-03 detector. The experimental results are of great importance for further development of the LVDS communication with radiation tolerance design which is intended for MAPS sensors in SOI technology.

2. SOI CMOS technology

The BOX layer isolates the CMOS electronics implemented on a thin, low resistivity, epitaxial silicon layer from a thick, high resistivity handle wafer. The sensitive element, reverse biased PN junction, is integrated in the handle wafer as illustrated in Fig. 1. In contrast to other SOI technologies, the thick film SOI provides a double well structure to shield the thin gate transistors from the BOX layer. The solution makes this technology a promising option for radiation imaging. This technology has been described in detail in [1,2].

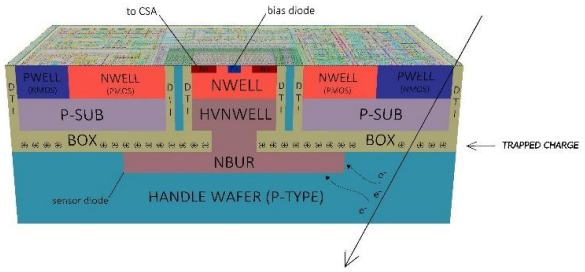


Fig. 1. Cross section and charge collection of X-CHIP-03.

3. LVDS drivers

The first prototype of the LVDS transmitter / receiver have been developed and manufactured in 150 nm technology. These blocks are standardized under ANSI/TIA/EIA-644, from [3], and completed only of 3.3V-transistors. The transmitter comprises of a current-mode driver which is not variable. In the transmitter design is a significant addition of the feedback [4].

A essential element of the receiver is differential amplifiers. The differential inputs are divided into two diff-

pairs of the diff-amps, and their output constitutes the output diff-pair of the output diff-amp. Current sources gates, which feed the diff-amps, are linked with NMOS current, electrons in NMOS transistors are faster as PMOS transistors. An accurate resistor with a value 100 Ohm is connected in parallel to the input due to the input current level and LVDS standard.

Specification of the LVDS block has been measured at the different frequencies with capacitance 0.1 pF, Fig. 2. At different frequencies has been sent over the transmitter (FPGA) 1024 bits long chain to the receiver (LVDS integrated circuit). The transmitter of the chip sends back the string into the FPGA, where the chains are compared.

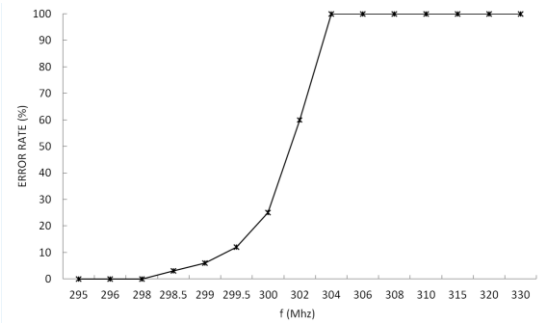


Fig. 2. LVDS drivers error rate, measured with diff.probe 0.1 pF.

The X-CHIP-03 includes the second prototype of the LVDS drivers. The schema of the second version form the schema of the first prototype, but the transistors dimensions and layout design have been changed due to the radiation resistance and strength of the drivers. The simulation of LVDS drivers is shown on Figure 3., where is the eye diagram of the differential signal at 0.5 Gbps.

The design of the LVDS IP blocks, as well as circuit simulation, laboratory measurement technology are described. The experimental results are of great importance for further development of the LVDS communication with radiation tolerance design which is intended for MAPS sensors in SOI technology.

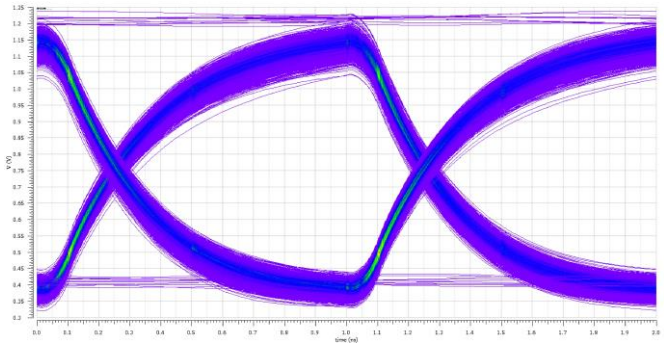


Fig. 3. Eye diagram of the differential signal at 0.5 Gbps.

Acknowledgements

Research described in the paper was supervised by Jiri Jakovenko, FEE CTU in Prague and Miroslav Havranek, FNSPE CTU and supported by the Technology Agency of the Czech Republic under project number TJ01000200.

References

[1] BENKA, T., et al., Characterization of pixel sensor designed in 180 nm SOI CMOS technology, *JINST*, 2018, vol. 13, no. 01., p. C01025.

[2] HAVRANEK, M., et al., 2018 MAPS sensor for radiation imaging designed in 180 nm SOI CMOS technology, *JINST*, 2018, vol. 13, no. 06, p. C06004.

[3] Electrical Characteristics of Low Voltage Differential Signaling (LVDS) interface circuits TIA/EIA-644-A, *Telecommunication Industry Association*, 2001.

[4] MANDAL, G., MANDAL, P., Low power LVDS transmitter with low common mode variation for 1GB/s-per pin operation, *2004 IEEE International Symposium on Circuits and Systems (IEEE Cat. No. 04CH37512)*, 2004, vol. 1, p. I-1120.

About Authors...

Tomas BENKA was born in Piestany, Slovak Republic in 1990. In 2014 has finished MSc at the Slovak University of Technology, Faculty of Materials Science and Technology in Trnava, Department of Informatics and Automation. Currently he studies doctoral degree at the Faculty of Electrical Engineering of Czech Technical University in Prague, Department of Microelectronics. His research activities include Pixel sensor designed in SOI CMOS technology.

Numerical optimization of coils system geometry for homogeneous magnetic field generation

David Novotný[1]

[1] Dept. of Measurement, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

novotd12@fel.cvut.cz

Abstract

Simple numerical method for designing of coil-system geometry and turn-ratio per section such leads to the high spatial homogeneity of magnetic field is presented in this paper. Presented approach can benefits from high computational power of contemporary PCs. Total magnetic field generated by multiple coil-sections, in any position relative to coil system is calculated with using Biot-Savart law and principle of superposition. Magnetic field distribution is then evaluated for its homogeneity area and compared to the previous result. Example of modeling 5-section coil, comparison to currently used coil-system are also presented.

Keywords

Magnetic field, homogeneous, Biot-Savart law, brute-force, optimization, geometry, Helmholtz, Lee-Whiting, coil

1. Introduction

Generation of a large area of homogenous magnetic field is useful or even necessary in many applications e.g. calibration of magnetometers [1], MRI, physical experiments. For such a purposes there are widely-known coil arrangements [2] (Helmholtz, Lee-Whiting, Merritt,..). All of them are pre-calculated based on analytical mathematics. With contemporary high computing performance of even ordinary home PC's, we are able to calculate these configurations using simply brute-force combinatorics and Biot-Savart law. Proposed method is similar to [3], but it is even simpler and thus easier to reconfigurate for desired purpose (e.g. generation of uniform magnetic gradient, generation of arbitary magnetic field with minimization of error,...).

From a Biot-Savart law can be derived equation 1 - magnetic field of current I loop with radius R out of symmetry axis - y. Such elliptic integral is solved numerically for each desired z, y position (drawing of coordinate system is in fig. 1).

Eq 1:

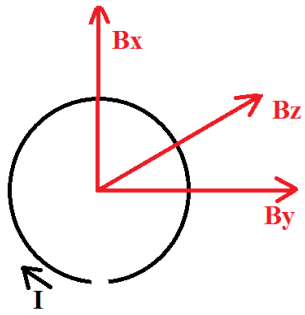


Fig 1: *Coordinate system of current loop*

2. Algorithm

Designed algorithm calculates magnetic field in one cut-plane for every possible coils system's configuration (with defined constraints and resolution).

At center position of coil system is considered maximum generated magnetic field. All calculated values of magnetic field are divided by such a maximum to get ratio. Area around center, where magnetic field ratio is higher than allowable deviation (e.g. 0,1%, 100ppm,..) is calculated. Symmetry of magnetic field along center in both axial and radial direction is used to speed up computation.

Continuous maximum searching is used to find coils configuration with largest area of homogeneity - latest candidate to the best configuration is saved.

For verification of computing correctness, two section coil system has been optimized by algorithm. Computed distance between sections leaded to known Helmholtz configuration (distance between section equal to radius of coils).

Simplified code of used algorithm for five-section coils is attached at the end of article.

3. Five-section coil optimization

For five sections coil arrangement, positions of each sections and turn ratios has been generated (table 1).

Section	Position	Number of turns
Coil 1	-1,1066·R	5.9·N
Coil 2	-0,3534·R	2.68·N
Coil 3	0	N
Coil 4	0,3534·R	2.68·N
Coil 5	1,1066·R	5.9·N

Table 1

Generated magnetic field at center can be calculated using equation 2 (with given current I, radius of coils R and N turns of center coil).

Eq 2: $B_z \cong 245 \cdot N \cdot R \cdot I \quad [\mu T, -, m, A]$

Comparison of homogeneity for used coil arrangements (Helmholtz, Lee-Whiting and simple solenoid) with designed five-section can be seen in fig 2. Spatial distribution of homogeneity is in fig 3. (2D - single cut-plane, x=0)

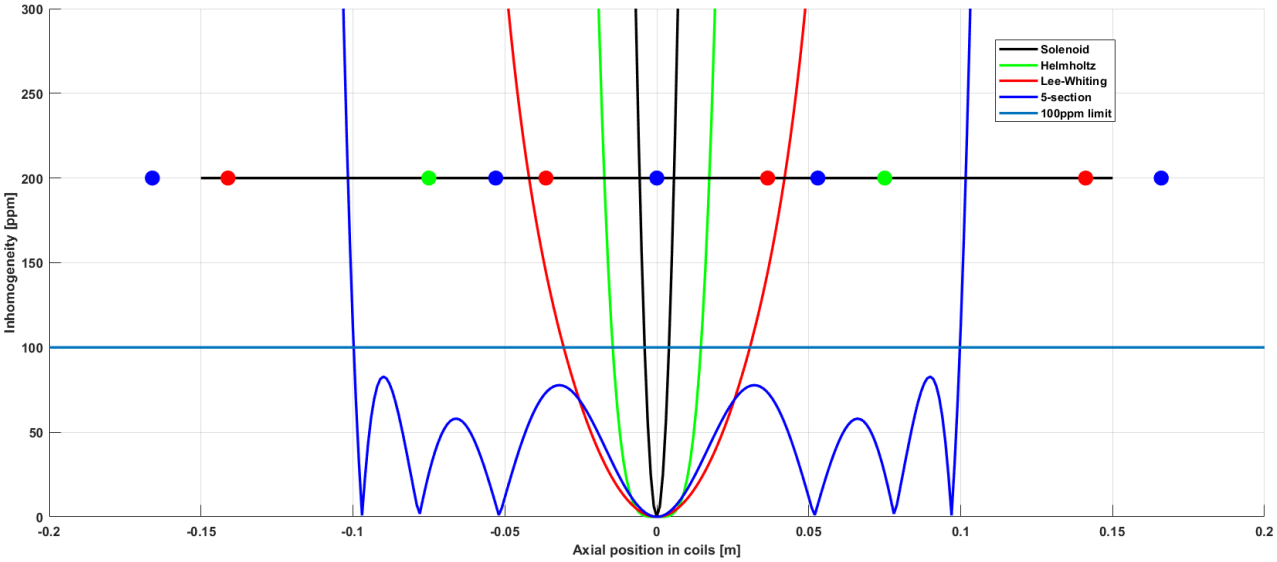


Fig 2: Comparison of coil system's inhomogeneity versus axial position (zero radial position) and coil section placement (solid line for solenoid) *(Simple solenoid, 2-section Helmholtz, 4-section Lee-Whiting, 5-section, 100ppm limit)*

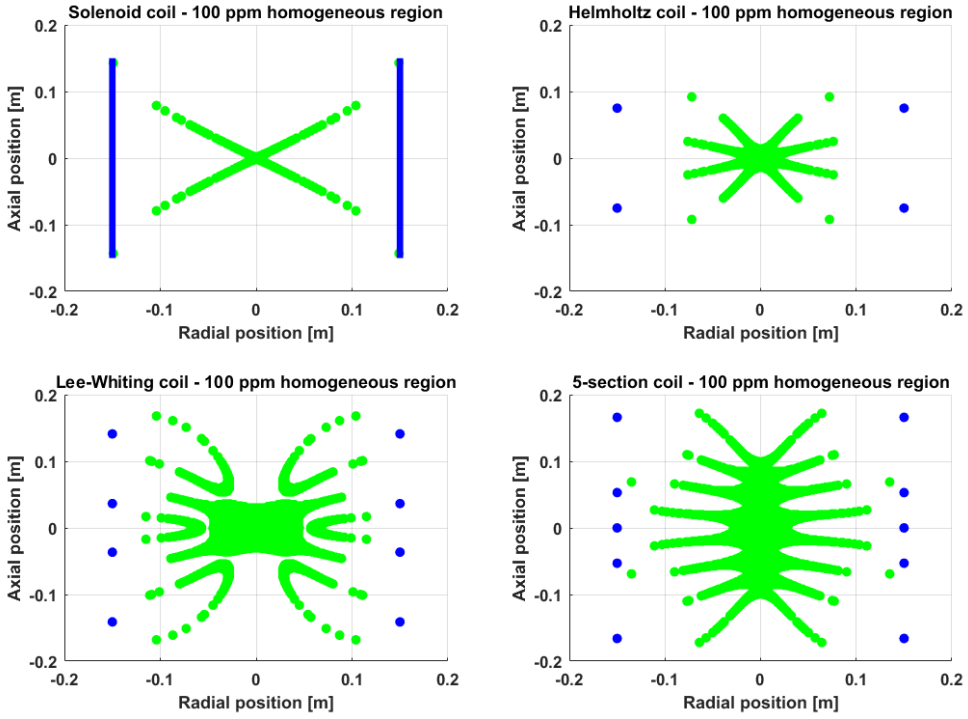


Fig 3: *Comparison of coil system homogeneity regions (2D cut-plane) (Simple solenoid, 2-section Helmholtz, 4-section Lee-Whiting, 5-section)*

4. Conclusion

Although complexity of brute-force algorithm is rising as fast as $O(c^N)$ with number of coil section, even with single core computation it is possible to calculate ideal 5-section coil configuration overnight. With using graphics card computation (thousands of computational cores with contrast to one CPU core) can be required time reduced to a few minutes or even orders of seconds. Besides simple configurations (the same radius for each section), more complex configurations can be optimized same way.

Novel coil configuration has been generated for 5 sections coil systems. In fig 2 is shown one-dimensional comparison of magnetic field inhomogeneity for simple solenoid coil, Helmholtz, Lee-Whiting and 5 section coils configuration. Two dimensional model of magnetic field homogeneity area (<100ppm inhomogeneity) is for comparison in fig. 3.

Acknowledgements

This project has been supported by the CTU research project SGS19/177/OHK3/3T/13. Design and evaluation of radiation tolerant magnetometer.

References

[1] M. Janosek, M. Dressler, V. Petrucha and A. Chirtsov, "Magnetic Calibration System With Interference Compensation," in IEEE Transactions on Magnetics, vol. 55, no. 1, pp. 1-4, Jan. 2019, Art no. 6000104.

[2] L. KIRSCHVINK, Joseph. Uniform Magnetic Fields and Double-Wrapped Coil Systems: Improved Techniques for the Design of Bioelectromagnetic Experiments. *Bioelectromagnetics*. 1992, (13), 401-411.

[3] Azpurua, Marco. (2012). A semi-analytical method for the design of coil-systems for homogeneous magnetostatic field generation. Progress In Electromagnetics Research B. 37. 171-189. 10.2528/PIERB11102606.

About Author

David NOVOTNÝ was born in Hořice, the Czech Republic, in 1993. In 2018 he obtained a master degree at Faculty of Electrical Engineering of Czech Technical University in Prague. After master degree studies he continues at doctor studies at the same faculty. His main research interest is magnetic measurement for space applications.

Simplified algortihm for 5-section coils (pseudocode):

```
% setting of parameters
R = 0.15; % Fixed radius of coil system
N1 = 100; % Fixed number of central coil turns
% algorithm settings
N3_min = N1;
N2_min = N1;
N3_max = 10*N1;
N2_max = 5*N1;
d1_min = 0;
d1_step = 0.001;
d1_max = R;
d2_min = 0;
d2_step = 0.001;
d2_max = R;
z_step = 0.001;
y_step = 0.001;

% algorithm
S_max = 0;
for N3 = N3_min:1:N3_max
 for N2 = N2_min:1:N2_max
  for d1 = d1_min:d1_step:d1_max
   for d2 = d2_min:d2_step:d2_max
     if d2>d1
      i = 0;
      % 2D mag. field computation
      for z = 0:z_step:d2
       j = 0;
    for y = 0:y_step:R
B(i,j)  =  contribution_of_coil(0,  N1, z, y);
B(i,j) +=2*contribution_of_coil(d1, N2, z, y);
B(i,j) +=2*contribution_of_coil(d2, N3, z, y);
       j++;
       end
       i++;
      end

S = calculate_homogeneity_region(B);

if S > S_max % maximum search
    S_max = S;
    best_values = [N3, N2, d1, d2];
end

     end
    end
   end
  end
 end
end

print(best_values);

%functions
function contribution_of_coil(d, N, z, y):
fi_step = pi/180; % 1 degree steps
% numerical integration
 for fi = 0:fi_step:(2*PI - fi_step)
  B=(N*(R-y*sin(fi)))/sqrt((R^2+y^2+(z-d)^2-
  2*y*R*sin(fi))^3);
 end
return B;

function calculate_homogeneity_region(B):
 B0 = B(1,1);
 S = 0;
 ppm_lim = 100; % 100ppm limit
 Berr = (B - B0)./B0;
 Berrppm = (Berr-1).*1e6;
 for i = 1:length(B(:,1))
  for j = 1:length(B(1,:))
   if Berrppm<ppm_lim
     S++;
   else
     break; % searching from center of coil
   end
  end
 end
return S;
```

Introduction of Photography Teaching at the Imperial and Royal Czech Technical University in Prague

Pavel Borkovec

Faculty of Electrical Engineering, Czech Technical University in Prague,
Jugoslávských partyzánů 1580/3, 160 00 Praha 6, Czech Republic

borkopa1@cvut.cz

Abstract. *The paper speaks about the introduction of Photography Teaching at the Imperial and Royal Czech Technical University in Prague. Professor Karel Kruis was the first full professor of photography and the paper propose to follow the birth of this discipline at the technical universities.*

Keywords

Photography, I&R Czech Technical University, Hermann Wilhelm Vogel, Josef Marie Eder, Karel Kruis.

1. Introduction

In this day and age, photography has more or less become part of everyone's life. Like every other invention, photography has also come a long way since its emergence at the beginning of the 19th century. Over the next few years, photography grew in popularity to such an extent that it became necessary to provide some scientific foundations. Its development from simple processes to complex but higher quality procedures and instruments hinged on the availability of well educated professionals, making it imperative to start offering photography classes at technical schools.

This article will focus on the emergence and stabilization of photography as a subject of study at technical schools. We will draw a comparison of its introduction at several technical schools in Germany and focus in more detail on its formal presentation at the Czech Technical University in Prague.

2. The invention of photography and its arrival in the Czech lands [1]

Visual arts were constantly refined with every generation until they reached their climax by the end of the 18th century. Numerous paintings and portraits possessed such quality that they were considered a faithful reproduction of the pictured reality. Their completion required both time-intensive work and extensive experience, as well as considerable skills and, most importantly, a great artistic talent. Under the circumstances of that time period, the creation of paintings and portraits also carried considerable financial costs. Therefore, it is surely not by chance that towards the end of the 1700s different ways of making these goods available to the wider public at lower costs and, most importantly, of entrusting their creation to "artists" who lacked the required talent possessed by great painters began to be explored. Moreover, the invention of lithography at the end of the 18th century led to further intensification of the efforts to seek new techniques of reproducing reality. At the same time, the level of chemical knowledge had reached a point where it became possible to put these ideas into practice. When peace returned to Europe after a prolonged period of Napoleonic wars, several men ventured in this direction and, independent of each other, succeeded in discovering the principle of photography in two countries, namely France and England. Some materials were found to be sensitive to light exposure, a quality which soon began to be exploited.

Fig. 1. Daguerreotype from 1837 by Louis Daguerre (*Reference: Louis Daguerre, Wikipedia*)

Three individuals are considered to be the fathers of photography: Frenchmen Niépce and Daguerre, who gave birth to the process known as daguerreotypy, and Talbot, an Englishman credited with a similar discovery at about

the same time whose technique became known as heliography. Not long after, Joseph Petzval, an Austrian physicist, also contributed to the initial invention of photographic principles with his outstanding calculations for the portrait lens. Once these foundations were laid and news of these inventions spread throughout Europe and America, photography entered a stage of rapid development. In addition to the plates on which the image was captured, it was necessary to improve the photographic device itself as well as the base materials on which the acquired images could be reproduced and preserved. This development can be traced all the way to the present day and includes numerous changes and major milestones which helped advance photography to ever higher levels and continuously simplify individual phases of the photographic process. Given the gradual development of photography and its application in everyday life, which was the case very soon after its invention, it was equally important to come up with new ways of improved and especially faster production. Following the manufactory period, when photographic technology was still in its earliest phases, the photographic industry emerged and began to develop at the end of the 19th century. As factory production is becoming mechanized, while standards and processes are being established, the entire sector enters its scientific phase.

Fig. 2. Talbot's heliography from 1840 by William Talbot (*Reference: William Talbot, Wikipedia*)

In the Czech lands, photography gained widespread recognition over a very short time period after its invention and enjoyed immense popularity. Even though its popularity gave rise to a considerable number of businesses and allowed for the establishment of a very dense photographic network, Czech lands had to rely on imported photographic materials up until the First World War due to the fact that the photographic industry in our country did not develop until the end of the century and Czech production was largely dependent on foreign manufacture and imports (initially from England and France and later on from Germany, after its industry developed).

The expansion of photography and its continual improvement and, above all, industrialization demanded that proper scientific foundations of this new discipline should be established. Whereas photography in the first half of the 19th century was able to get by with mostly homemade instruments and materials, in combination with improvements provided by gifted amateurs and photography enthusiasts, in the latter half of the 1800s further development was spearheaded by men of science and photography eventually began to make its way into academic laboratories and university lecture halls, while the first scholarly papers on the subject started to appear in scientific journals. Ever since its inception, photography has always been on the borderline between two disciplines: physics and chemistry. Physics was needed to formulate and calculate optical theories and to develop optimum methods for transmitting the visible reality through lenses and objectives, while chemistry was applied to the preparation of a wide variety of compounds, emulsions and agents capable of capturing, depicting and preserving this reality. And it was these two disciplines that photography began to emerge from in education.

Fig. 3. The earliest Czech daguerreotype by Bedřich Franz from 1841 (*Reference: Moravská galerie v Brně, Omnia.ie*)

3. Photography teaching at technical universities

3.1 Beginnings of photography teaching – German technical schools (Berlin, Vienna, Prague)

The favourable conditions and circumstances in the development of technical disciplines made it possible for photography to be integrated into the syllabus at technical schools during the 1870s.

The Royal Technical University of Berlin, where photography lectures were introduced in the mid-1870s, is

a good example of this. The classes were given by Professor **Hermann Wilhelm Vogel** (1834-1898) who lectured on photochemistry (a total of 2 classes per week with a focus on the chemical effects of light on simple and complex bodies, the principals of different types of photography, on the application of the chemical effects of light in technology and on photographic chemicals), on photography (2 classes per week with a focus on positive and negative processes, colour printing, procedures to be applied when photographing buildings, paintings, machines and models, including a description of photographic objectives) and supplemented his classes with two-hour practical training[2].

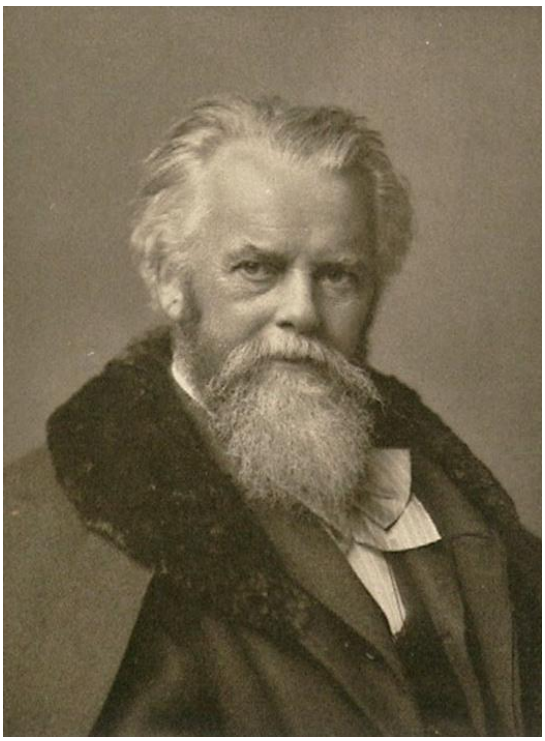


Fig. 4. Professor Hermann Wilhelm Vogel. (*Reference: Die chemischen Wirkungen des Lichts und die Photographie, Leipzig ,1883.*)

This was also the case for the Vienna University of Technology where photography and photochemistry began to be taught by **Josef Maria Eder** (1855-1944) in the mid-1870s. In the early 1880s, after his habilitation, Eder's research focused on photochemistry and practical photography. He received associate professorship in the academic year of 1892/93 and full professorship in 1903[3]. The teaching structure was not unlike that of Professor Vogel in Berlin. In his lectures, Eder explored the chemical effects of light and camera principles. He also expounded on the negative and positive procedures and explained the fundamentals of photo enlargement. During practical training sessions, he introduced his students to various cameras and to studio photography. He organized field trips to photograph buildings, interiors, machines, landscapes and rarely missed a good opportunity to take a snapshot. Eder also explored different methods of negative development and copy production[4].

Fig. 5. Professor Josef Maria Eder. (*Reference: Bildarchiv, ÖNB*)

Prague's technical schools responded to the development in education somewhat belatedly. The first school to offer photography classes was the Imperial and Royal Prague German Technical University. In 1893[5], Professor **Wilhelm Gintl** (1843-1908) established a photographic laboratory, paving the way for university study of photography in our country.

Fig. 6. Professor Wilhelm Gintl. (*Reference: Die k. k. deutsche technische Hochschule in Prag 1806-1906, Prague, 1906, p. 101.*)

3.2 Photography teaching at the Imperial and Royal Czech Technical University in Prague

The Imperial and Royal Czech Technical University in Prague fell behind other technical schools and began offering photography classes a number of years later in comparison to the above-mentioned institutions.

The first mention of the school's efforts to introduce lectures on photography dates back to 1896, that is 3 years after similar classes had been introduced at the Prague German Technical University. The proposal for the introduction of these classes was submitted by František Müller, Professor of Geodesy, at the faculty meeting held on 3 March 1896[6]. The faculty appointed a committee, composed of professors Müller, Preis and Domalíp, to discuss the possibility of introducing these lectures. A report was submitted to the faculty a month later and the introduction of photography lectures, including practical training in the extent of 3 classes in both semesters, was unreservedly recommended. Having unanimously approved the report, the faculty approached the ministry for the endorsement of classes on photography and photochemistry. The ministerial endorsement was not long in coming, primarily due to budgetary reasons. The faculty estimated the costs involved in the establishment of photography classes at approximately 10,000 florins. After some procrastination, the lectures were finally given a permission and a special office for photochemistry and photography was systematized in 1898. The faculty decided to appoint private senior lecturer **Karel Kruis** (1851-1917) to this office.

Karel Kruis had been employed by the university since the mid-1870s, firsts as an assistant and later on as an associate professor in organic dyes and distillery. His specialization in photography most likely stemmed from personal motives; however, it must have been inspired by professional pursuits as well. Being a chemist engaged in the processes of fermentation and yeast production, he was keen to explore various methods of capturing yeast cells using modern technology. He drew on his knowledge of photography to succeed in capturing microfilm images of yeast in 1896. This feat and his interest in photography helped him secure the post in the newly established discipline. Since Antonín Bělohoubek, Professor of Zymology, departed the university at about the same time, Karel Kruis was also put in charge of teaching zymology. This led to the unique merging of the two subjects, which can only be accredited to the personal interests and specialization of Professor Kruis (after his death, the teaching of both disciplines was separated since no close relationship could be found in this interconnection). However, thanks to this interconnection, the office of photography was elevated to a proper office and, in this respect, the Czech Technical University surpassed both the German Technical University in Prague and the Vienna University of Technology where a proper office was not established until three years later.

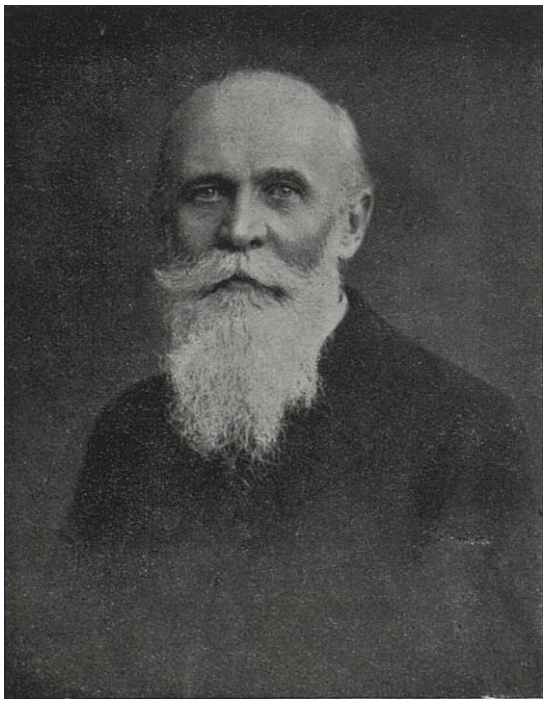


Fig. 7. Professor Karel Kruis. (*Reference: Český svět, 1918, n.19, p.8*)

The ministerial endorsement of the office of photochemistry and photography is dated as of 15 November 1899. Professor Kruis began giving lectures on photography immediately thereafter. In his classes, he taught about photography optics, objectives, and the camera and proceeded to explain the negative and positive processes. He would take pictures of various objects and demonstrated different ways of enlarging or reducing them. Moreover, Kruis lectured on microphotography and the required photographic equipment. He explained the methods of producing images of various preparations and pointed out the specifics of both the negative and positive micrographic processes. Hands-on training sessions were used to broaden the students' knowledge of the associated negative and positive processes as well as of enlargement and reduction procedures[7].

In drawing a direct comparison to the photography classes offered at German universities (see above), Kruis's personal interest in microphotography becomes evident. Whereas German universities did not devote any special attention to microphotography, Professor Kruis taught it to his students in great detail. Save for this difference, the teaching structure was essentially the same. All universities sought to introduce their students to the chemical processes of photography and photographic equipment. The principal aim of practical training sessions was to teach them take photographs of various buildings and objects as necessitated by industrial development and technical utilization.

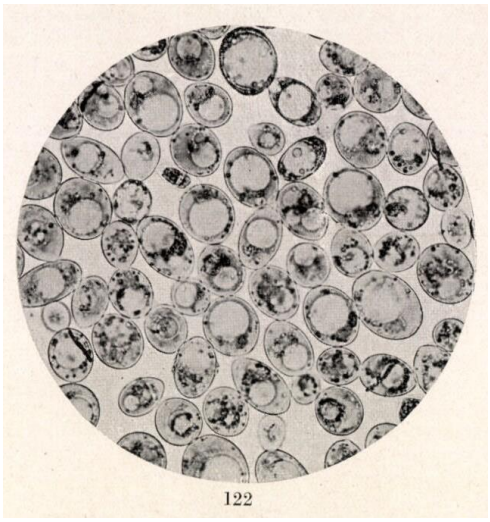


Fig. 8. Microfilm image by Karel Kruis. (*Reference: KRUIS, K. – ŠATAVA, J., O vývoji a klíčení spór, jakož i sexualitě kvasinek, Prague, Česká akademie, 1918, p. 83, n. 122*)

In conclusion, allow me to mention the condition of photography teaching at technical schools in Brno. At the Czech Technical University in Brno, lectures on photography were first presented in 1912. They were delivered by Professor **Vladimír Novák** (1869-1944). The introduction of these lectures is probably related to the establishment of the Technical Department, opened at the Czech Technical University in Brno in 1911. Their introduction certainly did not take long, as the considerable influence of the development in this new field and its inclusion in the curriculum of technical schools confirms. What is remarkable is that no similar lectures on photography were given at the German Technical University in Brno, which remained the only technical institution in the country that did not offer these classes.

Fig. 9. Professor Vladimír Novák. (*Reference: Provenio.net*)

4. Conclusion

Emerging at the beginning of the 19th century, photographic technology had undergone rapid development before it reached a phase in the second half of the 1800s where it became necessary to consider its introduction as a subject of study at technical schools. In the Central European region, it was incorporated into the curriculum in the 1870s. The early pioneers of this discipline were the technical universities in Berlin and Vienna. The Czech Technical University in Prague introduced photography teaching approximately two decades later. The first professor to teach photography there was Karel Kruis whose photographic knowledge and passionate interest in this new discipline had made up for the time lost with respect to photography teaching. It also contributed to the training of Czech professionals who gave birth to the Czech photographic industry several years later.

References

[1] According to BOUČEK, J., *Fotografická technika*. In: JÍLEK, F. et all., *Studie o technice v českých zemích*, Prague, Národní technické muzeum, vol. III, 1985, p. 215-238 and vol. IV, 1986, p 237-264.

[2] *Programm der Königlichen Gewerbe-Akademie zu Berlin für das Studien-Jahr 1878/79*, Berlin, 1878, p. 12.

[3] NEUWIRTH, J., *Die k. k. technische Hochschule in Wien 1815-1915*, Vienna, Selbstverlag der k. k. technischen Hochschule in Wien, 1915, p. 413-414.

[4] *Lektionskatolog, Studienpläne und Personalstand der technischen Hochschule in Wien für das Studienjahr 1903/1904*, Vienna, Verlag der k. k. technischen Hochschule, 1903, p. 20-21.

[5] STARK, F. – GINTL, W. – GRÜNWALD, A., *Die k. k. deutsche technische Hochschule in Prag 1806-1906*, Prague, 1906, p. 207.

[6] *Protokol sepsaný v sedění sboru professorského c. k. české vysoké školy technické v Praze, dne 3. března, 1896*, n. 26, fund Česká vysoká škola technická, n. 50, Protokoly, 1895/96, Archiv ČVUT, Prague, fol. 44.

[7] *Programm cís. král. České vysoké školy technické v Praze na studijní rok 1901-02*, Prague, Nákladem c. k. české vysoké školy technické, p. 66.

About Author

Pavel BORKOVEC is a Ph.D. student at study program *History of Technology* at the Czech Technical University in Prague.

Thundering the Engines of Flying Fortresses and Liberators over the "Palaba" Factory

Zbyněk NIKEL[1]

[1] Historical laboratory of electrical engineering of the FEE CTU in Prague, doctoral study program *History of Technology* (P 7108 / 7105V081), Czech Technical University, Technická 2, 166 27 Prague, Czech Republic

nikelzby@fel.cvut.cz

Abstract. *My paper deals with the system and organization of Air Raid Protection (ARP) in the firm Pála, j.s.c., during the Protectorate of Bohemia and Moravia. Managers of Palaba had to fulfill the dictation of the German nazi law on Air Defense (AD), darkening of buildings and vehicles. On the other hand, they had to face the consequences of possible raids by clusters of allied bomber aircraft or individual fighters. Palaba employees did not identify themselves with these measures and often avoided from the service in ARP patrols and fire fighter teams as well as medical (samaritan) guards.*

Keywords

Air Defense, Air Raid Protection, Bombing, Darkenig, Dogfights, extinguishing agent, Fire Fighter Team, Luftschutz, Luftwaffe...

1. Dogfights over the Protectorate

World War II raged. Although the Slaný town was in deep background nearly until the end of 1944, it was necessary to protect the factory, that, among other things, produced batteries for the warring Wehrmacht, against possible air raids.

My conference paper deals with the system of ARP service, and the elimination of possible air raid consequences, not at the strategic and tactic level of air warfare conducted by the British Air Force Marshal Arthur Harris: *„They sowed the wind, and now they are going to reap the whirlwind,"* but at the level of protection of property and employees of the firm Pála, j.s.c. For better orientation in this complex issue, it is briefly described the course of air war in the sky of the Protectorate of Bohemia and Moravia.

It can be stated that our territory was an important air battlefield, mainly on the turn of 1944 and 1945. This is related to the introduction of a long-range fighter aircraft P-51 Mustang to the operation service in the US Air Force. The Mustangs could protect the bombers throughout their mission, and thus, within air dogfights, still employ the Luftwaffe fighters engaged to the Air Defense system. That is why it was necessary to transfer many production capacities to the background, which was also the Protectorate until 1943. As soon as the Allies landed on the Apennine Peninsula and began breaking through to Italy, the situation changed. Austria and Czech Lands, occupied by German Nazis, were now threatened by air raids.

Basically, anywhere along the trajectory of attacking planes, an emergency release of bombs could cause a disaster. E.g. on March 25, 1945, there was*: „…stopped operation on railways Prague-Plzeň (section Zdice-Hořovice), Zdice-Protivín (Jince-Čenkov-Bratkovice and Březnice-Mirovice), Lysá nad Labem-Liběchov (Všetaty-Mělník) and Prague-Česká Třebová (Český Brod-Poříčany)."* Just this unintentional drop of bombs in Central Bohemia proves that Slaný could be hit at any time by lethal cargo from some stray airplane. And, apparently, just by coincidence of happy circumstances, at that time, the Factory for Electric Cells and Batteries in Slaný did not become the target of the Allied bombing effort. However, as it will be documented from archival materials, its near surrounding was bombed.

In conclusion of this chapter, it can be stated that the fighter component of Luftwaffe was almost defeated in the winter 1944–1945, and thus the Allied and Soviet Air Force could operate almost undisturbed over the Protectorate. It caused a lot of material damage to the Nazis, but, unfortunately, also to innocent civilians, and many people perished. The importance of the ARP of industrial buildings and private houses of the civilian citizens increased. Not because of the loyalty to the dying Nazi régime, but because of the German dictate to people, that they must become the members of a local, i.e. municipal and factory ARP system, or from a simple self-preservation instinct.

1.1 Air Raid Protection in the Third Reich

It should be noted, that the ARP system – "Luftschutz" – the Germans did not consider as a marginal, on the contrary, it was very sophisticated. For example, in the Cheb region, which belonged to the 3rd Reich:

- it was subordinated to Wehrmacht that equipped its particular stations,

- Schutzpolizei had responsibility for it, that was subordinated directly to Gestapo (Gehaime Staatspolizei),
- even local citizens were its members,
- in particular houses there was determined a Luftschutz member, whose duty was to check:
- darkening (verdunkelung),
- extinguishing agents,
- order in the protective shelter and its equipment,
- during air raid, he was responsible for preventing the panic and in time hidening the house citizens.

The members of Luftschutz were regularly trained and had special powers. The paper proves on the specific example of the firm Palaba that the Reich Law on Air Defense (Luftschutzgesetz) from June 26, 1935 (Reich Code I., p. 827) was entirely applied to even in the central bohemian royal city Slaný during WWII.

Fig. 1. Jaroslav Jan Pála boarding a car that has shielded headlights.[1] Keeping the darkening was generally mandatory.

2. Description of the Air Raid Protection System in the firm Pála, j. s. c.

In the 3rd Reich, and therefore even in the Protectorate Bohemia and Moravia, there was the Air Raid Protection prescribed by law – Reichsluftschutzgesetz. Its latest valid edition, obligatory also for the Palaba employees, was promulgated in August 1943.

This law divided Air Raid Protection to:

- **direct protection of every individual**, i.e. self-protection, carried out in houses,
- **extended protection** that included public buildings, restaurants, hospitals, etc.,
- **factory protection**, applied in enterprises, industrial plants, etc.

Integral, and indeed essential, parts of the ARP were darkening at night and concealment at daytime.[2]

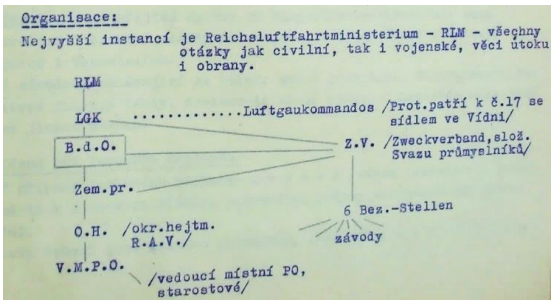


Fig. 2. In the picture above there is a brief diagram of the Air Raid Protection at the 3rd Reich level.[3]

Notbelegschaft (Necessary Staff, NS) had to remain in place in the case of an air raid (engineers in boiler rooms, etc.)

Einsatzgruppe (Operational Group, OG) consisted of firefighters, paramedics, decontamination teams, veterinary teams and drivers.

Bereitschaftsgruppe (Readiness Group, RG) served to cover OG losses or to change a member in OG and had the same organization as the OG.

In the industrial plants, that were operated on shifts, there had to be both OG, and the RG on each shift.

Commander (WLS-Leiter) was not allowed to leave his place and had to be reachable to all ARP components. The crew consisted of his deputy, telephone operators, messengers. Warnkomando (Early Warning Command) received reports from Flugmeldedienststellen (Air Reporting Service) and gave warning to other Air Defense components: military troops, public authorities and Institutes, as well as Warnstellen (Warning Stations). Telephone exchanges were set up in industrial plants, so called Unterwarnstellen (Subordinate Warning Stations), that worked if the plant received an air raid hazard report from another factory.

In the event of an air raid alarm, the work in a factory did not have to be interrupted if the factory patrols were watching the movement of allied, i.e. "enemy" aircraft.

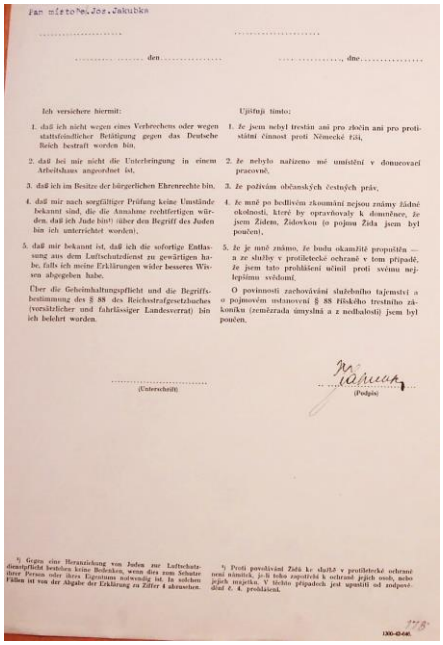
For this decision – to interrupt the work or not – Werkführer (owner or manager of the plant, work-leader) was responsible.

Everyone couldn't serve in the factory ARP. A special category was the Jews, who were to be called into service in the ARP if it was *"needed to protect their people and property"*. Moreover, the calling for duty of Jews to the ARP was governed by special provisions. So, the Jews could not actually serve in the ARP. See the document copy below.

[1] Undated photo from an archive of the company *Bateria Slaný CZ, Ltd.*, from the period of the Protectorate of Bohemia and Moravia.

[2] Essentially, it was the layout and modification of the various buildings so as they visually integrated with their surroundings.

[3] State Regional Archive (SOA) Prague, fund Pála, a. s., Slaný, NAD 1255, Registered Unit Number: 57. *Závodní protiletecká ochrana*, No. 141.



Ich versichere hiermit:

1. daß ich nicht wegen eines Verbrechens oder wegen staatsfeindlicher Betätigung gegen das Deutsche Reich bestraft worden bin,
2. daß bei mir nicht die Unterbringung in einem Arbeitshaus angeordnet ist,
3. daß ich im Besitze der bürgerlichen Ehrenrechte bin,
4. daß mir nach sorgfältiger Prüfung keine Umstände bekannt sind, die die Annahme rechtfertigen würden, daß ich Jude bin[1]) (über den Begriff des Juden bin ich unterrichtet worden),
5. daß mir bekannt ist, daß ich die sofortige Entlassung aus dem Luftschutzdienst zu gewärtigen habe, falls ich meine Erklärungen wider besseres Wissen abgegeben habe.

Über die Geheimhaltungspflicht und die Begriffsbestimmung des § 88 des Reichsstrafgesetzbuches (versätzlicher und fahrlässiger Landesverrat) bin ich belehrt worden.

Čestně prohlašuji tímto:

1. Že jsem nebyl trestán ani pro zločin ani pro protistátní činnost proti Německé říši,
2. že nebylo nařízeno mé umístění v donucovací pracovně,
3. že požívám občanských čestných práv,
4. že mně po bedlivém zkoumání nejsou známy žádné okolnosti, které by opravňovaly k domněnce, že jsem Židem, Židovkou (o pojmu Žida jsem byl poučen),
5. že je mně známo, že budu okamžitě propuštěn — a ze služby v protiletecké ochraně v tom případě, že jsem tato prohlášení učinil proti svému nejlepšímu vědomí.

O povinnosti zachovávání služebního tajemství a o pojmovém určení § 88 Říšského trestního zákoníku (zrádzrada úmyslná a z nedbalosti) jsem byl poučen.

[1]) Gegen eine Heranziehung von Juden zur Luftschutzdienstpflicht bestehen keine Bedenken, wenn dies zum Schutze ihrer Person oder ihres Eigentums notwendig ist. In solchen Fällen ist von der Abgabe der Erklärung zu Ziffer 4 abzusehen.

[1]) Proti povolávání Židů ke službě v protiletecké ochraně není námitek, je-li toho zapotřebí k ochraně jejich osob, nebo jejich majetku. V těchto případech jest upustiti od zodpovězení čl. 4. prohlášení.

Fig. 3. Every member of the Palaba ARP had to sign a declaration of integrity, and also that he was not a Jew.[4]

The structure described above was essentially copied by the Palaba ARP system, which was further developed in the individual buildings of the factory to the smallest details.

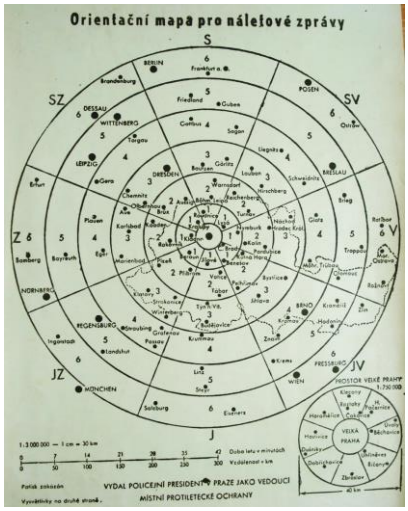
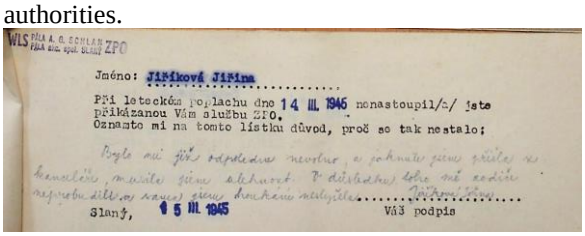


Fig. 4. A flying map, more precisely a "map for air raid reports", issued by the head of local ARP, i.e. the police president in Prague, showing the borders of the Protectorate of Bohemia and Moravia at a scale of 1 : 3,000,000; that is, 1 cm on the map corresponds to 30-km distance as the crow flies. The center of concentric circles is located in Prague. Circle radii are used to easily subtract the distance of the aircraft viewed from the point of interest. At the bottom right there is a "large Prague area" at a scale of 1 : 750,000, i.e. the radius of the outer circle represents a distance of 20 km.[5]

Sequence No.	Appointment	Name
1.	**Firm manager and the ARP Head:**	Director Rudolf J. Pála
	I[st] Deputy:	Director Antonín Hnátek
	II[nd] Deputy:	Vice Director Josef J. Jakubka
2.	**The I[rst] Shelter Head:**	Jindřich Nassler
	Deputy:	Otakar Pála
3.	**The II[nd] Shelter Head:**	Karel Šip
	Deputy:	Václav Janovský
4.	**The III[rd] Shelter Head:**	Jaroslav Solfinger
	Deputy:	Ing. C. S. Kašpar
5.	**Riot Service Head:**	Vice Director Josef J. Jakubka
	Deputy:	Dipl. Eng. C. S. Kašpar
6.	**Fire Fighting Service Head:**	Technological Manager Václav Pála
	Deputy:	Rudolf Mrština
7.	**Health Service Head:**	
	The I[st] Health Service Team Head:	Albína Stádníková,
	Deputy:	Helena Düllková
	The II[nd] Health Service Team Head:	Helena Pajerová
	Deputy:	M. Rovnanová-Kejlová
	The III[rd] Health Service Team Head:	Olga Fišerová
	Deputy:	Jaromír Šefčík
8.	**Telephone Service Head:**	Božena Opeltová
	Deputy:	Květa Charvátová
9.	**Signal Service Head:**	Vladimír Rákos
	Deputy:	Ladislav Skála

Tab. 1. The names of Heads and their Deputies of the Air Raid Protection organization of the Palaba factory in July 1942.

The ARP system in Palaba seems to be perfect at first sight, both in terms of organizational, and materially-technical security. The Palaba employees, involved in the ARP system, should have even been financially rewarded. But the reality was far from the desired ideal. People did not properly perform their duties towards the factory and the Reich within the Protectorate of Bohemia and Moravia. The factory ARP manager regularly wrote records of absences and misconducts of persons determined for services in the Air Raid Protection system during training alarms or when real dangerous appeared in the sky. It is hard to say to what extent the ARP Head thought seriously about his reprisals, or if it was only the alibistic reports that were supposed to protect him from the wrath of the Nazi authorities.



Jméno: Jiřiková Jiřina

Při leteckém poplachu dne 14. III. 1945 nenastoupil/a/ jsem přikázanou Vám službu ZPO.
Oznamuji mi na tomto lístku důvod, proč se tak nestalo:

Slaný, 15. III. 1945 Váš podpis

Fig. 5. The apology of the Palaba employee Jiřina Jiříková for her absence from the ARP service in March 1945. The Palaba employees sometimes slept hard, sometimes they were sick or hit by sudden nausea and often did not have enough courage to face the danger lurking in the sky over the Slaný town.[6]

Employees involved in the ARP were trained in the skills of firefighters, paramedics, messengers, etc., according to sophisticated manuals. Under alarms, they had to wear ARP labelled sleeves on their arms, but, above all,

[4] State Regional Archive (SOA) Prague, fund Pála, a. s., Slaný, NAD 1255, Registered Unit Number: 57. *Závodní protiletecká ochrana*, No. 178.

[5] State Regional Archive (SOA) Prague, fund Pála, a. s., Slaný, NAD 1255, Registered Unit Number: 57. *Závodní protiletecká ochrana*, No. 81.

[6] State Regional Archive (SOA) Prague, fund Pála, a. s., Slaný, NAD 1255, Registered Unit Number: 57. *Závodní protiletecká ochrana*, No. 117.

they had to serve essentially as soldiers in addition to hard work at the factory. They should have been, and perhaps have been, rewarded, or they should have received a compensation for fare or clothing that could be damaged during an ARP intervention.

If someone did not take part in daily or night-time services, he had to apologize or prove his apology by medical confirmation.

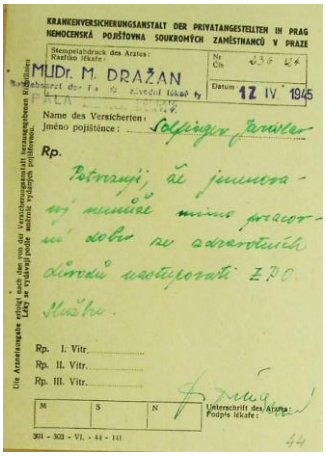

KRANKENVERSICHERUNGSANSTALT DER PRIVATANGESTELLTEN IN PRAG
NĚMECKÁ POJIŠŤOVNA SOUKROMÝCH ZAMĚSTNANCŮ V PRAZE
MUDr. M. DRAŽAN
Datum 12 IV 1945
Name des Versicherten: Jméno pojištěnce: Solfinger Jaroslav
Rp.
Potvrzuji, že jmenovaný musíval mimo pracovní dobu ve zdravotních důvodů nastupovati Z.P.O.
Rp. I. Vitr.
Rp. II. Vitr.
Rp. III. Vitr.
Unterschrift des Arztes: Podpis lékaře:

Fig. 6. Confirmation of the physician Dražan, M.D., of the firm Pála, j. s. c., that the insured Jaroslav Solfinger cannot serve in factory Air Raid Protection for health reasons.[7]

As mentioned above, the most dangerous period in terms of air raids was for Central Bohemia, and thus also for the Slaný town, on the turn of 1944 and 1945. The frequency of "visits" to Allied aircraft, and thus the activation of the Palaba ARP, is evidenced by the selected records of announced alarms at the factory. It is important to realise that the sirens were roaring day and night. The Palaba employees had to be in permanent stress and often did not sleep.

"December 16, 1944: Alarm sounded by sirens at 11.45. Ended at: 13.40 – Overflights of enemy aircraft. Dropping two big and a few small bombs at 800m distance from the factory building.

February 14, 1945: Daily alarm sounded by sirens in town at 12.20 and ended at 13.30. Overflights of aircraft. Night alarm sounded by sirens at 20.40, ended 22.12.

February 15, 1945: Night alarm sounded by sirens at 00.20, ended at 01.30; overflights of aircraft. Daily alarm sounded by sirens at 11.07 – ended at 15.35. Overflights of aircraft over factory buildings.

March 2, 1945: Alarm sounded by sirens in town at 10.16 and ended at 12.20. In the vicinity of the Slaný town, near the God's grave (u Božího hrobu), 4-engine enemy aircraft was shot down."[8]

This state escalated after the allied bombing of Prague on February 14, 1945, when the management, and especially the officials of the Palaba ARP, i.e. Rudolf Pála and Dipl. Eng. Kašpar, did not know what to do with the increasingly dissatisfied employees, supported by the factory committee, who no longer intended to risking their lives, basically for the dying Nazi empire.

Its coma came on Saturday, May 5, 1945, when the Prague (Czech) uprising broke out. Jaroslav Jan Pála was arrested on Sunday, May 6, 1945. The end of the renowned factory for producing of electrochemical cells and batteries Palaba in royal city Slaný was approaching.

Acknowledgements

Research described in the paper was worked out in the frame of the 23rd International Student Conference on Electrical Engineering, POSTER 2019, under the supervision of Marcela Efmertová and Jan Mikeš (Historical laboratory of electrical engineering, Faculty of Electrical Engineering of the CTU in Prague) and within the grant No. **SGS17/131/OHK5/2T/13**: *Technical Professions: Technical Engineers and the Czech Society (Between Scientific Milieu, National Prestige and Entrepreneurial Activities) in the Years 1881–1945*.

References

[1] State Regional Archive (SOA) Prague, fund Pála, a. s., Slaný, NAD 1255, Registered Unit Number: 57. *Závodní protiletecká ochrana.*

[2] KÁRNÍK, Z. *Malé dějiny československé (1867–1939).* Ed. 1. Prague: Dokořán, 2008. 502 p., [24] p. Imagery Attachment. ISBN 978-80-7363-146-8.

[3] PLAVEC, M. *Útoky spojeneckých hloubkařů u Podlešína na konci dubna 1945*. In BARTONÍČEK, P., ed., ČEČRDLE, J., ed. and HRABÁNKOVÁ, B., ed. Slaný a Slánsko v letech 1939 – 1945: Proceedings from the Historical Conference, Slaný 2010: November 6, 2010 in the Grand Municipal Centre. Slaný: The Museum of National History in Slaný in cooperation with the Slaný Municipal Authority, ©2011. 167 p. The History of Region Slaný in the 20th Century; Vol. 5. ISBN 978-80-904657-2-5., p. 134–143.

[4] PLAVEC, M. *Bomby pod Řípem: nálety na Kralupy nad Vltavou, Neratovice, Veltrusy, Hněvice, Mělník, Roudnice [i.e. Roudnici] nad Labem a další sídla během druhé světové války*. Ed. 1. Cheb: Svět křídel, 2008. 278 p. Svět křídel; 104. ISBN 978-80-86808-56-7.

[5] PLAVEC, M., VOJTÁŠEK, F. and KAŠŠÁK, P. *Praha v plamenech: nálety na hlavní město za druhé světové války.* Ed. 1. Cheb: Svět křídel, 2008. 423 p. Svět křídel; 105. ISBN 978-80-86808-58-1.

[6] VÁLKA, Z. *17.12.'44: největší letecká bitva nad Protektorátem.* Olomouc: Votobia, 2004. 137 p., xvi s. Color Imagery Attachment. ISBN 80-7220-200-6.

[7] VESELÝ, M. *Hvězdy nad Krušnohořím: letecká válka nad severozápadními Čechami 1944 – 1945.* Ed. 1. Prague: Naše vojsko, 2005. 126 p., [16] p. Imagery Attachment. ISBN 80-206-0793-5.

[8] MATĚJÍČEK, L. *Chebská křídla v souvislostech: příběh historicky prvního vojenského letiště v českých zemích a nacistické letecké továrny.* Ed. 1. Cheb: Svět křídel, 2013. 463 p. ISBN 978-80-87567-35-7.

[7] State Regional Archive (SOA) Prague, fund Pála, a. s., Slaný, NAD 1255, Registered Unit Number: 57. *Závodní protiletecká ochrana*, No. 44.

[8] State Regional Archive (SOA) Prague, fund Pála, a. s., Slaný, NAD 1255, Registered Unit Number: 57. *Závodní protiletecká ochrana*, No. 13–16.

[9] RAJLICH, J., SEHNAL, J. *Konec Luftwaffe v Čechách.* Part 1. Plastic kits revue, 1992, Year 2, No. 11, p. 4–7.

[10] National archive Prague, fund Ministerstvo železnic I, pasteboard No. 4544, Inventory No. II-21092-45, II-21093-45, II-221094-45 and II-21095-45.

[11] *K výročí největšího bombardování Prahy.* Prague City Bulletin 12, II (XLIX), 1946, p. 174.

About Author…

Zbyněk NIKEL was born on July 4, 1966, in Bílovec. He is an employee of the Historical laboratory of electrical engineering, Faculty of Electrical Engineering of the Czech Technical University (CTU) in Prague, and student of the all-school doctoral study program *History of Technology* (P 7108 / 7105V081). Nikel's studies concerning the planned dissertation with the working title *History of the company Pála et al., factory of electric cells and batteries in Slaný, 1889 (1919)–1945*, was approved by the Scientific Council of the CTU doctoral study program *History of Technology* on February 8, 2018.

The Harvester SM-500 – the machine, which brought joy and sorrow

Martin VYCHYTIL[1]

[1] Czech Technical University, Jugoslávských partyzánů 3, 160 00 Prague 6, Czech Republic

vychyma1@cvut.cz

Abstract. *The contribution deals with the history of the emergence and development of the combine, the SM-500. In its time this was a very reliable and powerful machine, which brought several design upgrades and not between harvesters competition. First is outlined the tradition of harvesting whack-jobs in Czechoslovakia, which is divided into 3 stages (1928-1945, 1945-1950 and from the year 1951). They are given the politico-economic reasons that influenced the introduction of combine harvesters in Czechoslovakia. Due to these reasons and circumstances, the need arises for the executive of the mower. The enhancement involved the company Agrostroj Prostějov, which began to mass-produce combine harvester ŽM-330, and then began to develop a harvester employ the SM-500. They are described the conditions in which the harvester originated, explained the name of the SM-500, and justified need for the introduction of the harvester. There is also the importance of the harvester from the point of view of the technical. Technical data harvester SM-500 are compared with the technical parameters of the combine harvester Fortschritt E-512, which is in Czechoslovakia, the most used and the SM-500 replaced, because it was ordered for political reasons the development of the harvester SM-500 stop, which meant for design engineers of Agrostroj a big disappointment, which captured in their memories. The Text is supplemented with many photographs, there prove the timelessness of the machine SM-500. The importance of the article lies in capturing the history of the machine, which had large technical and structural assumptions, however, to him there is no single comprehensive publication.*

Keywords

Agrostroj Prostějov; Comecon; Fortschritt; Harvester; the SM-500

1. Tradition of combine harvesters in Czechoslovakia

The history of harvesting whack-jobs on the territory of our state is very diverse. Recently, however, become the subject of research not historians, but experts with the technical (or mechanical) education.

The Production machines, which would reap cereals in the fields, began in 1928. At that time, the first articles on combine harvesters began in Czechoslovakia. Not only these articles, but also various exhibitions, they informed the Czechoslovak public with the harvest of grain. The first harvesting mutilator could in Czechoslovakia to see the visitors to the Spring economic exhibition in Prague. It was a machine of American manufacture. The greatest impact on our harvesting machines, however, was Germany, where the research was harvesting whack-jobs the biggest. This development was however interrupted by the Second world war.

In the years 1945-1950 were then harvesting machinery to Czechoslovakia imported under the us humanitarian aid to the UNRRA (United Nations Relief and Rehabilitation Administration), which was formed on the basis of the Washington agreement of 9. November 1943 for the purpose of economic aid to states damaged for 2nd world war, which did not have own resources to the recovery of agriculture. In the case of agriculture it was about the importation of livestock, raw materials and machinery for the production of industrial fertilisers, but also agricultural machinery including combine harvesters. At the same time, the ministry of agriculture carried out the first detailed research harvesting whack-jobs, mainly from the standpoint of economy of operation compared to the current way of harvesting.

From 1951 to combine harvesters are getting quickly into the awareness of not only professional agricultural public, but also to the laity. The Harvesters at that time were already imported in large quantities, were the first comparative trials harvesting whack-jobs and created the first agro-technical standards of the harvest. In developing the design of the machinery should be of great significance testing on the field and in the laboratory.

2. Need executive machine

The Need to develop a powerful harvesting machine is associated with the politico-economic regulations and situations. After February 1948, when there was a seizure

of power by the Communist party of Czechoslovakia and the subsequent collectivization, they had to peasants through the limits and following the example of his greatest partner – the Union of Soviet Socialist Republics (USSR), start to grow wheat on large fields in the greatest quantity. To be able grain harvest as efficiently as possible, they needed the combine that would be powerful. In 1949 began to compulsorily arise of the unified agricultural cooperatives, which marked one of the few ways the state forms of the agricultural economy, which lasted until 1990. In addition to them, work also machine tractor stations. They owned at the beginning of and during the 50's years the vast majority supplied by the whack-jobs. Most of the harvesting whack-jobs transferred to the collective farm after 1959.

2.1 Harvesters before the SM-500

Before the emergence of the combine, the SM-500 were Czechoslovak farmers dependent on the import of harvesters from abroad. The machines were either imported in bulk from Eastern Europe and Southeast European countries (S-4, SK-4, Fortschritt E-512) or less in western Europe (Laverda, Claas, Sampo, and others.

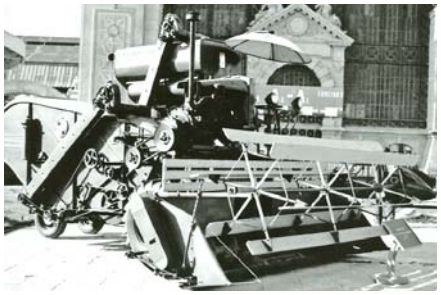


Fig. 1. Harvester S-4. [5]

The first harvester threshers, which is on our fields began to appear, came from the USSR. These were S-3 and above all S-4.

2.2 Harvester S-4

It was the first self-propelled combine harvester produced in the USSR. The design was based on the American machines International Harvester No. 123SP. The development of the project was supported by M. A. Pustygin and I. S. Ivanov, in the Institute of Applied Science and Building Science for Construction of Economic Machines. It was serially produced in 1947-1958 in Taganrog and later in Krasnoyarsk.

Fig. 2. Mikhail Andreevich Pustygin (1906- 2012). [15]

The harvester was 4300 mm wide, 3600 mm high, 7000 mm long, and weighed 3800 kg. **The Engine** (gasoline six-cylinder ZIS-5K) with an output of 39 kW / 53 k, was stored on the platform next to the driver, **the width of the mower** was 4 m.

It had a four-bar **shaker** with a length of 2680 mm and a width of 900 mm. The maximum **throughput** was 2.4 kg/s. The **tray volume** was 1700 l. The **maximum transport speed** was 14 km/h.

The biggest **drawback** was the frequent clogging of the threshing the influence of a longer grain and the high power consumption of the motor (17-30 l/ha).

2.3 The Stalinec 6

Stalin's concept was based on drawn American harvesters. It was founded in Rostselmash in Rostov on Don under the direction of I. F. Vasilenkov, A. V. Krasnichenko, P. N. Proshumin and I. F. Popov.

Fig. 3. Ivan Fomich Vasilenko (1894-1980). [15]

The **engine** was a four-cylinder petrol engine with a power output of 29.4 kW / 40 hp. The **mowing width** was 4.88 m. The **tray volume** was 1830 l.

The disadvantage of the machine was laborious handling and poor control and high operating costs.

Fig. 4. Stalinec 6. [5]

2.4 Harvester SK-3

This harvester was formed under the leadership of designer Isakson in the GSKB (государственное специальное конструкторское бюро) in Taganrog and later at Rostselmash in Rostov on Don.

Fig. 5. Harvester SK-3. [5]

The machine was imported into Czechoslovakia between 1961 and 1963. The **mower width** was 3.2 m.

2.5 Harvester SK-4

In the enterprise Rostselmash began to produce a modernized type SK-4. Together with the SK-3, machines used to be used in modern times for very modern harvesters, which have been imported into nearly 40 countries.

Fig. 6. Harvester SK-4. [5]

2.6 Hungarian threshers

In addition to the combine harvesters imported from the USSR, the machines imported from Hungary appeared to a lesser extent. It was all about harvesters type AC 400, ACD 343 and LM 330, which originated in the enterprise EMAG. These threshing machines, however, often followed the soviet machine (S-4).

The disadvantage of all previous machines, it was also a very uncomfortable working environment for the driver. The harvesters did not have a glazed cabin and the drivers had to face a very dusty environment. In addition, the combine engine was often built-in at the driver's head so the machine was very noisy. Combined with low performance, which has not reached higher average yields, wet climates and higher grain purity, a solution has been sought. One of them came to the firm Agrostroj Prostějov.

3. The experience of Agrostroj Prostějov with the construction of combine harvesters

Many of the harvesters mentioned in Chapter 2 had an impact on the construction of combine harvesters at Agrostroj Prostějov. A great inspiration was the S-4 harvester for the local designers, whose Soviet backgrounds were rewritten. The 2 prototypes of the ŽMS-4 were created in 1954. From the Soviet pattern it differed with a mowing bar with a continuous worm, a chaff hood suction pipe drawn into a trailer, and another cover for the end of the shaker. The couplings, gearboxes and steering were used from Czechoslovak trucks. Tests and test results of the machine served to the construction of the ŽM 330.

It was created 3 prototypes ŽM 330. The machine was powered by a **diesel engine** Tatra T114, but it was produced with the Tatra T 924 (44.1 kW / 60 hp). The width of the mower was 3 m. **The novelty** was a double grain cleaning (a classic 3-sieve mesh cabinet and a small mesh cabinet with 2 sockets to which hot air was blown from engine cooling to help drain grain), worm emptying the hopper, more space over the sifter, mower adapter with hydro pneumatic suspension and self-copying of the soil surface in the horizontal direction. **The throughput** was 3.3 kg/s, **the maximum output** was 3.7 kg/s. **The transport speed** was 15 km / h.

The deficiency of the machine, however, was a heavy weight (5440 kg), which was partly reduced to 4200 kg.

The success of the Prostějov engineers was certainly the fact that the machine had the best results on the international benchmarking tests in the GDR.

The transfer of production of ŽM 330 was interested in Bulgaria, Poland and Hungary, where production was transferred at the end of January 1957.

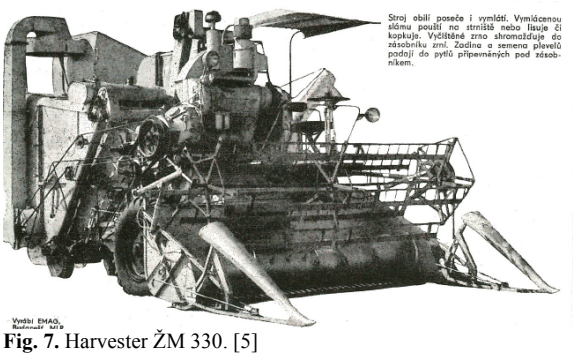


Fig. 7. Harvester ŽM 330. [5]

3.1 Conditions of formation of the harvester SM-500

The company Agrostroj Prostejov did not want to stay only for copying Soviet threshers. For his employees was a great prestige and excitement not only from the local but also from the international point of view to develop a harvester that would be in decisive technical parameters a few steps ahead of all the hitherto produced and imported combines. The time certainly has helped the emergence of the harvester SM-500. The first half of the 1960s was a period of political relaxation, which also gave rise to major projects.

3.2 Name of the SM-500

It was then the custom of naming the harvesters with the first letters of their names, for example СК-4 (from the Russian trademark “самоходный комбайн,

4-я модель"), ŽM 330 (from the Czech trademark "žací mlátička") with the width of the cutter bar of 330 mm.

The letters SM express the combine harvester (but it is also the theory that they express the slope - the harvester was designed on slopes). In some archive sources, even the name of a combine harvester with higher slope access appears. It is certain that number 500 represents the swath of the cutter bar.

3.3 Large slopes, smaller yields

The higher slope accessibility of the combine was one of the many reasons for its introduction. At the time of the harvester SM-500 for the climbing ability area in cereals considered the terrain with a slope greater than 8–10°. While mechanical harvesting of cereals in planar areas with a predominant slope of the land up to 8-10° was satisfactorily solved by a conventional harrow, harvesting of cereals in gradients above 8-10° was not satisfactorily solved. The use of conventional combine harvesters (SK-4 or E-512) for large harvesting was paid for by a high increase in product losses (often 10-15%), which was particularly serious in breach of occupational safety. The solution was neither due to economic reasons (high production costs influenced by machine complexity and relatively low series) nor special combine harvesters with a throughput of around 2 kg/s, which were commonly used in mountainous areas of Western Europe and mountain combs with automatic positioning.

During the development of the new combine harvester, it was also necessary to respect the specific natural conditions of the slopes of Czechoslovakia (higher water precipitation, higher slope of the land, rugged form of plots, frequent occurrence of stones, poor condition of the access roads, different structure of vegetation and its higher ingenuity, lower yields per hectare and demanding safety requirements on slopes).

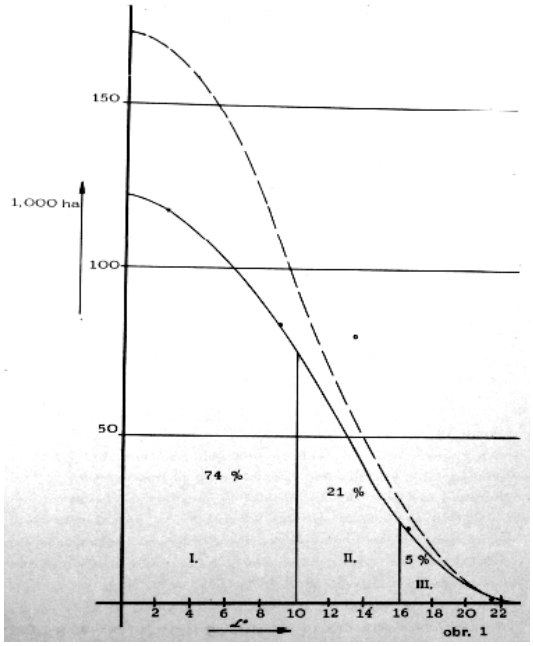


Fig. 8. Representation of crops on land according to the slope.[14]

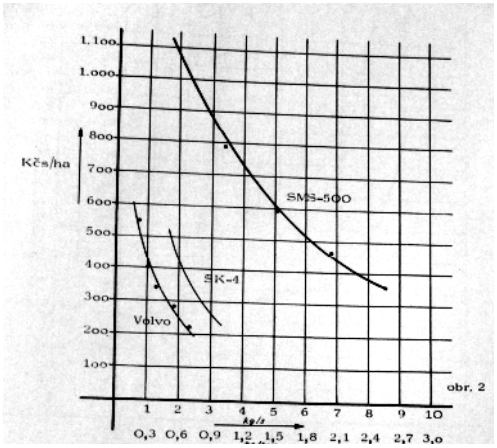


Fig. 9. Dependence of operating costs on mowing and peeling of cereals in terms of operating conditions. [14]

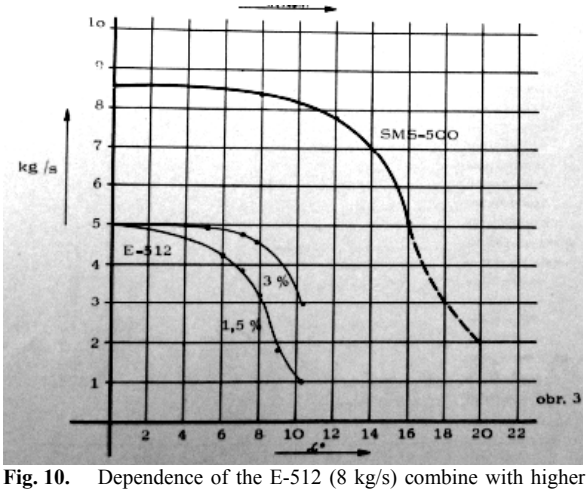


Fig. 10. Dependence of the E-512 (8 kg/s) combine with higher slope availability compared to SM-500. [14]

3.4 Solution in SM-500

The solution should be the SM-500 combine with slope accessibility 16°. This was to be achieved by a low centre of gravity. The grain container was located at the bottom of the machine in the space between the drive and the steering axle and the engine placed on the drive axle. The location of the stack at the bottom near the centre of gravity has created a favourable situation in that the stability of the machine was virtually independent of the stack filling level.

3.5 Technical parameters and advantages of the SM-500

The SM-500 combine harvester was designed to harvest seed crops under conditions of Czechoslovak agricultural large-scale production, from cereals to grasses, clover, legumes, maize and oil crops. That was the answer to the technical solution.

The harvester was designed as a kit of basic functional groups, some of which like the drive axle with hydraulic control, motor unit and operator workplace are

common to self-propelled forage harvester, harvesting sugar beets and potatoes, or on a smaller scale with a heavy tractor and farm truck.

The machine was a mowing table with a detachable mowing roller that was stored for transport by a hydraulic conveyor. Then from your own threshing machines, the chassis, the motor unit and the platform operator, on which was placed the cabin. Its solutions were based on the most demanding requirements of safety, hygiene and work culture. Thresher skeleton formed steel frame. The cutting table was equipped with a device for controlling the pressure on the ground from the maximum to the zero, the mowing machine was equipped with transverse copying, so it automatically monitored the terrain configuration in its entirety. The mowing device was equipped with a built-in gearbox, which allowed adapting the conditions of the harvesting of the harvested crop. The relatively small wheelbase with hydraulic control and hydraulic drive ensure maximum mobility and operability in rugged sloping terrain and dead ends, minimal unproductive times.

Machine dimensions in mm

	Working	Transporting
Length	8,950	12,450
Height	3,450	3,450
Width	6 190	2,980

Machine weight in kg

Dry	Curb
7 990	8 360

Cutter bar in mm

Basic	Administrative
5 000	6 000

Engine power: 180 hp

Throughput (in wheat at a yield of 50 q / ha at a grain ratio of straw of 1: 1): 8-10 kg

Slope accessibility: 10°

Surface yield in direct harvest: ÷ 2 ha per hour

Area power at split harvest : ÷ 2,5 ha per hour

Seasonal performance: 300-500 ha per hour

Attachments: collecting device, crusher and spreader of straw, adapter for the harvest of corn to grain, adapter for the harvest of clover.

3.6 Advantages and novelties of the SM-500 combine

Among the advantages of the SM-500 combine was the large area output; high quality of work; high culture, safety and hygiene of work; advanced compact design using modern technology; easy operation and maintenance.

From other combines the SM-500 differed by hydrostatic travel, hydrostatic control, dust extraction directly from the rail and hydraulically tilting chimney. These elements were unprecedented at the time because they did not apply to the combine harvesters. The difference was also in throughput. While the other combines were around 4-5 kg/s, the SM-500 should have been 10 kg/s already. It follows that the SM 500 was a lot better than the harvesters used until then.

In its time, the SM-500 was one of the first threshers to have six shredders and a thawing canal of 1.5 m long. In addition, the mowing line with hydraulic leveling has not been a long time for harvesters by standard. The machine also excelled in its ease of use. All controls had a driver at hand, only 1 lever was used to drive the machine forward and backward. The harvester went forward with the lever, the harvester stopped with the reverse lever without the clutch, and the harvester was driving backward on the lever. There was no other farm machine at the time.

The SM-500 also looked nice for its time. It was created at a time when, besides technical claims, it also placed emphasis on design. This trend moved to Czechoslovakia from the west in 1968-69. The SM-500 excelled in graceful colors, and at that time the favorite angular shapes.

3.7 Failure and sadness

During 1969, voices began to be heard that international agreements should also be observed in the development of the SM-500 combine harvester. Great attention has been paid by Fortschritt from the GDR. German engineers were very interested in the technical solution of the machine and its appearance. This, in turn, did not appeal to the workers of Agrostroj Prostějov, who took the matter as a rivalry. This fact was also reflected in the first name of the future combine harvester - RIVAL.

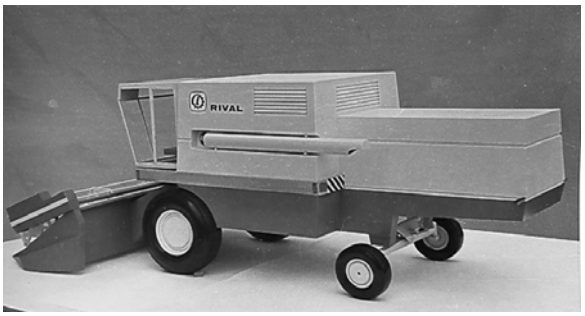


Fig. 11. Model SM-500 combines still with RIVAL inscription. [18]

The report contains a record of different types of tests (e. g. economic, technical). *"At the same time, the Fortschritt E-512 combine was launched from* the NDR*, which was the state-of-the-art machine used by us, but the SM-500 surpassed almost all of its parameters"*[18] says Vladimír Michálek of the Museum of Agricultural Technology.

The failure could CMEA spotting pressures that favoured some states and determine what will be built where.

The saddest ideas of the creators of the new combine began to be fulfilled. The last hope for the rescue of their project was the visit of former President Ludwig Svoboda to Agrostroj Prostějov. They believed that he is the man who might have plans of the workers from Prostějov CMEA to push through. Even the visit of President Ludvík Svoboda did not succeed. The Rules of Comecon was clearly given and so in 1972 the Czechoslovak government decided to Agrostroj Prostejov development of a new combine harvester ended.


Zvláštní vydání

STRÁŽ LIDU

VZÁCNÁ NÁVŠTĚVA V DEN HISTORICKÝCH VÝROČÍ ■ 14. ŘÍJNA 1918 GENERÁLNÍ STÁVKA A VYHLÁŠENÍ SOCIALISTICKÉ REPUBLIKY ■ 15. ŘÍJNA 1920 PRVNÍ ČÍSLO STRÁŽE LIDU ■ 14. ŘÍJNA 1970 PRESIDENT REPUBLIKY V NAŠEM MĚSTĚ ■ SRDEČNÉ OVACE ■ ČESTNÝ OBČAN NAŠEHO MĚSTA NAVŠTÍVIL RADNICI ■ CHLÉB SE SOLÍ Z ÚRODNÉ HANÉ PRVNÍMU HOSPODÁŘI

SLYŠEL JSEM MNOHO DOBRÉHO O VAŠÍ PRÁCI

Hanácká obec vítá presidenta

Fig. 12. Visit of President Ludvík Svoboda in Agrostroj Prostějov. [12]

3.8 Memories of those who have been involved in the development

At first, great enthusiasm, but in the end, a huge disappointment can be found in the words of some of the workers involved in creating the SM-500 combine.

Designer František Suchánek (he was in charge of the front part - mower adapter, loading conveyor, longitudinal and transverse ground copying)

"As far as work is concerned, we did overtime and we spent a lot of time in the trials because it was tested seven days a week, from morning to evening, and when the machine was driving, we went from 6am to 10pm and when the machine did not go - there was some defect, so he was remodelling during the night. "He adds," The enthusiasm was great until it was rumoured to go down. "

Designer Vladimír Špaček

"They envied us that we were sitting without dust in the cabin, envious that we had a higher performance (almost double)."

Designer Stanislav Brácha

"When the engine was started, when the first prototype was finished in the workshop and you heard the thunder of the engine, it was an experience that a bit resembles the experience of an athlete who will achieve a certain performance that only dreamed."

"There was a certain amount of regret and hopelessness at that time, because we realized that there was no such opportunity here, that such a machine would never evolve here and do it. And when we looked back at everything we were willing to do, there was regret, maybe anger, and especially a sense of frustrated opportunity. Maybe it could be likened to Mount Everest mountaineering - you've practically everything you've done before, done, just waiting for the exit and now you get the message that you have to pack the suitcases and go back down and back and not only that you will never get to that one."

4. Fate and present of the SM-500

The SM-500 combine harvester never went into mass production. There were only a few prototypes. At present, only two prototypes have been preserved. One prototype is exhibited at the Museum of Agricultural Technology in Čáslav and the other in the Slovak Agricultural Museum in Nitra. The harvester was also seen in 2016 at the show "Země živitelka" in České Budějovice.

Fig. 13. Harvester SM-500. [19]

The exact number of prototypes originated and the place where they are located is not exactly known. According to information from multiple sources, the SM-500 served as a model in the construction of the Don 1500 in the USSR and the Fortschritt harvesters in the GDR. The Fortschritt E-516 harvester has a similar look. The Fortschritt E-512, E-514 and E-516 harvesters

prevailed in the crop in Czechoslovakia, but in many technical parameters the SM-500 lagged behind.

Fig. 14. Cab of the SM-500. [20]

Fig. 15. Harvester SM-500. [21]

4.1 The SM-500 compared to its successor

Overview of important technical data of the E-512 combine harvester:

Length of machine in mm	
in the transport position	12,470
in the working position	8,120
Width of machine in mm	
in the transport position	2,893
in the working position	6,943
Height of the machine in mm	3,820
Weight of the machine in kg	
with the header	6,880
without the header	5,900
Working width of the header in mm	4,200
Slope accessibility in %	18 ÷ 20
Engine power in hp	105
Clearness in kg/s	5

In comparison, it is obvious that the SM-500 had a wider shot of the cutter bar, increased throughput and higher engine power. In addition, the SM-500 harvester could boast more shakes (E-512 had 4) or automatically tilted chimney.

Fig. 16. Fortschritt E-512. [16]

Fig. 17. Fortschtitt E-516. [17]

Acknowledgements

Reserch described in the paper was supervised by Ing. J. Mikes, Ph.D., FEE CTU in Prague and supported by the Student Grant Competition of CTU under grant SGS17/131/OHK5/2T/13 *Technical Professions: Technical Engineers and the Czech Society (Between Scientific Milieu, National Prestige and Entrepreneurial Activities) in the Years 1881-1945.*

References

[1] BERNHARD, K., PEHAL, F. and collective. *Sklízecí mlátička E-512.* 2st ed. Praha : Státní zemědělské nakladatelství, 1974, p. 24 – 27.

[2] MALÉŘ, J. *Modernizace sklízecí mlátičky SK 4.* Zpravodaj zemědělské techniky, technické provozní a obchodní zprávy Ústředního podniku zemědělské techniky. Brno, Agrozet 1971, vol. 5, No. 12, p. 202 – 204.

[3] MICHAL, V. *Žací mlátičky u nás: Dobrý pomocník v nedostatku sil.* Zemědělské družstevní noviny. 1947, vol. 2, No. 36, p. 2.

[4] NOVOTNÝ, F. *Kapitoly z historie techniky pro sklizeň obilnin ve světě (10): Historie zavádění přímé sklizně obilnin v bývalém Sovětském svazu a v českých zemích.* Farmář, informační měsíčník pro zemědělce. 2001, vol. 7, No. 5, p. 92 – 94.

[5] STEHNO, L. and collective. *Historie sklízecích mlátiček.* 1st ed. Praha : Profi Press, 2014, p. 14 – 15, 18, 25 – 30, 33 - 34.

[6] VÁLKA, J. *Porovnání parametrů sklízecích mlátiček nasazených v ČSSR.* Zpravodaj zemědělské techniky, technické provozní a obchodní zprávy Ústředního podniku zemědělské techniky. Brno : Agrozet 1989, vol. 23, No. 6, p. 75.

[7] VRÁNA, V., ZAHRADNÍČEK, J. *Československá žací mlátička ŽM-330.* Za socialistické zemědělství, časopis pro šíření vědeckých poznatků do zemědělské praxe, 1955, vol. 5, p. 1 – 40.

[8] Samochodné žací mlátičky dodané z UNRA (Moravský zemský archiv in Brno, fond Agrostroj Prostějov).

[9] RVHP – připomínky k návrhu unifikované řady SM (motory) (Moravský zemský archiv in Brno, fond Agrostroj Prostějov).

[10] Maďarsko – písemnosti k předání výrobní dokumentace mlátičky ŽM-330 do MLR RVHP (Moravský zemský archiv in Brno, fond Agrostroj Prostějov).

[11] Zpráva o možnosti výroby sovětských kombajnů SK-4 (Moravský zemský archiv in Brno, fond Agrostroj Prostějov).

[12] Stráž lidu – zvláštní vydání, vyšlo v Prostějově v den návštěvy prezidenta republiky Ludvíka Svobody 14. 10. 1970.

[13] Sklízecí mlátička SM-500, prospekt (Moravský zemský archiv in Brno, fond Agrostroj Prostějov).

[14] Návrh postupu vývoje a zahájení výroby sklízecí mlátočky s vyšší svahovou dostupností (Moravský zemský archiv in Brno, fond Agrostroj Prostějov).

[15] Википедия: *свободная энциклопедия.* [online]. Accessible from: ru.wikipedia.org

[16] *Fortschritt E 512, Das waren Zeiten*. Pinterest. [online]. Accessible from: https://www.pinterest.de/pin/692287773944970152/

[17] *Mähdrescher Fortschritt E 516 B - MDW*. Pinterest. [online]. Accessible from: https://cz.pinterest.com/pin/700239442032758435/

[18] *Cycle "Zašlapané projekty": Combine harvester for museum* (in the Czech original *"Kombajn pro muzeum"*). Czech Television. [online]. Accessible from: https://www.ceskatelevize.cz/porady/10209988352-zaslapane-projekty/409235100061016-kombajn-pro-muzeum/

[19] *Pradědečkův traktor.* FarmWEB. [online]. Copyright© Accessible from: https://www.farmweb.cz/index.php?page=view_image&id=NDAwWDc4OTFYODM5NzY=#skok

[20] *Pradědečkův traktor.* FarmWEB. [online]. Copyright© Accessible from: https://www.farmweb.cz/index.php?page=view_image&id=NDAwWDc4OTFYODM5ODE=#skok

[21] *Muzeum zemědělské techniky.* FarmWEB. [online]. Copyright© Accessible from: https://www.farmweb.cz/index.php?page=view_image&id=MjM1MFgxNjI4MlgxNzgyNDk=

About Author

Martin VYCHYTIL

was born in 1988 in Ústí nad Orlicí. In the years 2008 to 2011, he graduated from UHK – PdF in bachelor study program History and Russian language with focusing on education and he got bachelor's degree. Subsequent master's field "Teaching for secondary schools – History and Russian language and literature" he graduated from the same university in the years 2011 to 2013. He got Mgr. degree. In 2017 he graduated from the Charles University – Faculty of Physical Education and Sport. Here he got PhDr. degree in the program Physical Education and Sport. He is a combined doctoral student of the All-university doctoral study program History of technology at The CTU in Prague since October 2016. The topic of his dissertation is "History of OEZ Letohrad in the years 1945 to 1967". Currently, he also teaches history, Russian language and law at the Business Academy Dušní in Prague.

History of shoe-making in Skuteč up to 1949

František KREJČÍ[1]

[1] Institute of History of the Philosophical Faculty of the University of Hradec Králové, nám. Svobody 331, 500 02 Hradec Králové, Czech Republic

frantisek.krejci.1@uhk.cz

Annotation. *The subject-matter of the paper titled "History of shoe manufacturing in Skuteč up to 1949" is analysis of suitable archive materials and professional regional literature with objective to characterize the history of this industrial sector using the example of the small town of Skuteč. The town of Skuteč, which lies in the foothills of the Iron Mountains and the Bohemian-Moravian Highlands is known from the viewpoint of economic history thanks to two facts. Firstly, this region was rich in the deposits of various minerals, which also included marble that was hewed here from quarries already at the turn of the 60s and 70s of the 19th century and is still hewn to a limited up to the present. Secondly, a very strong shoe manufacturing tradition is tied to the town itself and its oldest written records date back to 1534.*

In terms of the thematic and time aspects, the paper shall be divided into three sections. The first section shall on the basis of the knowledge from professional literature on the history of the shoe industry from the earliest written records up to 1939. The second part will focus on the reconstruction of selected shoe manufacturing companies from their establishment to their nationalisation and the third part will describe the methods used to manufacture shoes up to the 1st half of the 20th century.

Keywords

Shoe industry; Skuteč; economic history

1. History of shoe manufacturing in Skuteč and its reflection in regional literature from the first written records up to 1939

1.1 Beginning of the manufacturing of shoes up to the abolition of guilds

Early written records of shoe manufacturing in the Skuteč region exist already **from the 15th century**. The shoes subsequently became a trading article at various fair markets in the vicinity. At the time, the shoemakers came together and established the Guild of St. Cryspin. [1] The history of the Skuteč shoe industry thus dates back to the turn of the Middle Ages and Early New Age, on which we get information from the so-called **Shoemakers' New Order**, which was issued in 1534 by **Jindřich Šťastný of Wallenstein and at Rychmburk**. This Order clearly defined the distribution of the work among the shoemakers. The right to make new shoes was wielded by the so-called new shoemakers, while the old worn shoes were repaired by the "cobblers". From the outset, **shoes were made in households** and were sold at the anniversary markets in Bohemia and Moravia.

In the era of the **Government of Maria Theresa**, **military boots** were mainly made in the Skuteč region and supplied to the **garrisons of the forts of Josefov, Hradec Králové** and also supplied to **Brno**. It was necessary to equip the **Habsburg Monarchy** soldiers who participated in the battles for the Austrian heritage with quality military boots. During the Napoleon wars, the **Guild Masters** supervised the production of the military shoes and had the right to punish the guild members for bad workmanship. The orders were thus made directly from Vienna to ensure that production continued and the ordered shows were delivered in time. After **1830** the Guild Masters were replaced by the **military suppliers** - lifranti (derived from the German word *liefern* - to supply). Up to **1848** the orders for the military forces were fulfilled irregularly and the resulting products were characterised by poor quality. This period did not close until the issue of the public declarations for supply of shoes to the military forces in the same year and **in 1853** the **activities of the guilds were finally abolished officially.** [2]

1.2 Establishment of the first shoe-making factories and trades up to the end of 1918

In **1859 the new trades licensing regulations were published**, which predetermined the extensive development of the shoe industry. The home made shoes lost significance and the increase of factory production came quite quickly. The abolition of the guilds gave rise to the opportunity of **capitalist enterprise**, which started developing in the town of Skuteč from about the year 1860. **David Borges** can be tagged as the first entrepreneur in the shoe industry.

His further entrepreneurial activities also include the management of trade in cloth, a grocery and a liquor production plant. Borges's sons - Mořic and Josef did not rank among the shoe manufacturers, but had adequate funds to pay the shoe masters. Mořic Borges even owned two shoe-making workshops at Palacké náměstí no. 379 and 380. Apart from the Borges family, other Jewish families in Skuteč also made a living from the shoe industry, **for example, Pollak, Schlesinger, Falk.**

.

Fig. 1. View of Mořic Borges's house, where he operated his shoe-making trade. (The Municipal Museum of Skuteč)

This moment marked the beginning of the development of the **factory production of shoes** in Skuteč, which was started by the aforementioned Borges family and which was continued by the local masters Rakušan, Šlemr, Prášek and others. **Jan Prášek**, who founded a shoe factory jointly with **Vincenc Rakušan** can be termed the **first Skuteč factory owners**. Their cooperation later failed and each of them founded an own factory. [3]

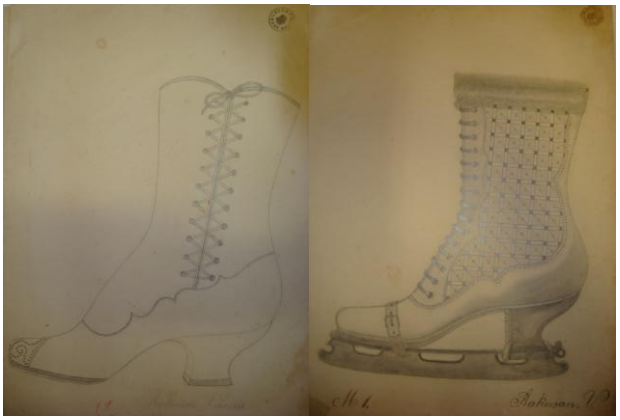


Fig. 2. Shoe drawing from the shoe-making shop of master Vincenc Rakušan. (The Municipal Museum of Skuteč)

The increase in the establishment of the shoe-making shops and factories was **enormous** in Skuteč **from the 60s of the 19th century up to 1914** and this is supported by the contemporary records. Prior to the First World War, it is possible to talk about the operations of about **19 shoe manufacturers** who employed almost **750 workers**. The established companies thus included the enterprises **Rosner and Mandelík**, the **Karel and Josef Kusí** brothers or the enterprise of **Václav Gregor** whose trade was taken over by **Josef Kunc**. Also worth mentioning is the enterprise of **Václav Bláha**, which was established just before the outbreak of the war. Up to **1918** the major markets for the Skuteč shoes apart from the domestic towns were also the countries of former Austria-Hungary, which changed after the declaration of the First Czechoslovak Republic. [4]

1.3 The fate of the Skuteč shoemakers after the declaration of the independent Czechoslovak Republic up to the eve of the Second World War

After the collapse of Austria-Hungary the Skuteč shoemakers encountered major **market problems** since after the declaration of the independent State of Czechoslovakia **customs duty was imposed on the shoe exports**, which was gradually reflected in their rising prices. For this reason, it was necessary to change the sales strategy and **focus more on the domestic market**. However, strong competitor was formed in this market in the form of the **Baťa Works**. The coincidence of these factors resulted in the fact that many of the shoe-making masters were compelled to abandon their trade. Certain livening up of the economy occurred **between 1919 to 1921** when new trades were established and factories were built, which manufactured shoes of various types, or were narrowly specialised. The specialized factories focused mainly **on the manufacture of outdoor shoes, working and children's shoes**.

The good time for the developing shoe enterprises of the Skuteč shoemakers however reached its peak after the entry of the already mentioned **BAŤA** company of Zlín into the trading sector. **The machine-made shoes from the** Baťa Works were **cheaper** and the company had **better distribution** options among the consumers. The influence of this company on the Skuteč Region is again reflected in the contemporary statistics. **From 1928** it is possible to see a declining trend **both in the number of functional enterprises** and in the number of **employees** who often sought work in **some local open pit mine** after discharge from the factory. [5]

A clear deviation from the declining trend in the number of manufacturers and employees was the following year, i.e. **1929** when **560 workers** worked **for 18 manufacturers**. Specifically in **1931**, there were **9 manufacturers** in Škuteč who employed an average of **136 workers**. For comparison, **three years earlier** prior to the outbreak of the **Great Economic Recession,** the number of manufacturers had almost doubled and the number of employees in the factories was up to four-fold **(17 manufacturers and 508 workers)**. An example of the shoe-making companies that were established in the inter-war period, but did not survive the difficult period of the economic recession, not even after the establishment of specific business brands, it is possible to state Jan Novotný's enterprise **NOVEA** or the enterprise of the shoe-making masters Václav Habal and Bohuslav Bláha known as **TILLIA**.

In this difficult period, the **shoe-making masters, factory owners and entrepreneurs in the branch** who started to specialise in the **manufacture of luxury shoes** had the highest chance of survival, which cannot be said of the rank and file shoemakers, who were often compelled to opt only for repair of old shoes. **The period 1929 to 1931/1932** is in the sign of extensive bankruptcy proceedings, which apart from some less successful factory owners also afflicted the smaller manufacturers. The outcome of these events was the convention of a **protest meeting of the shoemakers of Eastern Bohemia**, which took place on **8 February 1931** with the participation of representatives from **Pardubice**, **Chrudim**, **Holice**, **Litomyšl**, **Havlíčkův Brod** and other towns. However, the earlier boom of the Skuteč shoe industry was not restarted by the **military orders for the Czechoslovak Army** prior to the outbreak of the Second World War. [6]

2. Selected shoe-making companies and their development up to 1949

If we ignore the Skuteč small-scale shoe manufacturer whose entrepreneurial activities mostly did not survive the world economic recession, during the 1st half of the 20th century several personalities from the ranks of the shoe masters who succeeded in the commercial market and whose business operations were often terminated during the post-war nationalisation after 1945.

The following treatise shall be just about these successful entrepreneurial subjects whose trade was unique in a given sense.

2.1 The Rybička brothers - Shoe manufacture in Skuteč

Establishment of the company of the Rybička brothers falls under the year **1909**, whereas their joint factory went down in the history of Skuteč **as the first machine production of shoes in Skuteč**. The founder brothers - **Karel and Josef** - each specialised in a different type of livelihood. Karel was more interlinked with the company because he worked as a master shoemaker while Josef was the co-owner of the real estate property. From **1923** the company was transformed into a **public trading company** and the first problems with shoe sales gradually appeared. In 1930, on the basis of competitive pressures, this company went bankrupt and Josef moved to Pardubice to join his brother Karel, who had moved there two years previously and established an own enterprise. Their production plant was taken over by the company **Unitas (Goodyear), s. r. o.** and later taken over by **Jaromír Brzokoupil**, whose enterprise was nationalised in 1948. [7]

Fig. 3. Shoe factory of the Rybička brothers.
(The Municipal Museum of Skuteč)

2.2 František Hošpes - Shoe factory in Skuteč

The enterprise of **František Hošpes** ranked **among the most successful enterprises in the region**, which is supported by the size of the company capital, whose portion was used to fund the construction of the new factory building from **1924**. The establishment of the company dates back to the same year when it was established by cooperation of other three shoemakers, namely: **Čeněk Chour, František and Josef Štěpánek**. Thanks to its establishment on co-ownership basis, this company also survived the **period of the economic recession**, although **František Štěpánek left its management in 1931**.

It is possible to state that in the inter-war period this company could be proud of a relatively extensive customer base and solid prosperity, which is supported by the list of customers not only from the **Czechoslovak Republic**, but also from **Denmark**, **Germany** and the **USA**. The history of this company was not curtailed even by the German occupation, which cannot be stated in relation to the post-war nationalisation in 1948, which afflicted it and was subsequently integrated in 1949 into the **national enterprise SVIT**. [8]

Fig. 4. Advertising leaflet of František Hošpes's company. (The Municipal Museum of Skuteč)

3. Shoe-making technology up the 1st half of the 20th century

Just as applies to other **crafts related with the processing of leather**, shoe production from the outset took place mainly in **the household conditions of the shoemakers**. Every shoemaker owned the following aids: **a pairing-knife** (shoemaker's knife for cutting leather), chisels, metal pricker for making holes, shoetrees, **shoemaker's stirrup** (leather stirrup use to fix the shoe in process to the knee) and the "**shoemaker's table**" - working table with three-legged stool. The shoemaker needed a source of light during his work and for this reason he often did his work in front of his dwelling. During the 19th century the shoemakers used the **kerosene lamp** to light their working space in the households. The **average shoemaker** made a **pair of shoes every 2 days**, but many a time it took him even several days to make more demanding shoes.

The exclusive position in the shoemaker hierarchy was held by the **master**, who had several helpers around him - these were the **journeymen and apprentices**. It was in the competency of the master to purchase raw materials, allocate work, receive orders and sale of goods on the market. **Although the journeyman was a trained shoemaker** who however still lacked adequate experience after the examination to be able to perform the trade. He acquired the necessary experience during the **so-called journey, which lasted 3 to 5 years**.

The apprentices had the **worst position** in the shoemaker's workshop, which is supported by how they were treated. Mostly, this concerned a **boy aged about 10**, who also performed helping chores in the household. **Hand-making of shoes** prevailed up to **the first half of the 19th century**. A significant milestone was therefore the implementation of the **sewing machine**, which was generally an integral part of the normal equipment of a shoemaker's workshop **in Bohemia at the beginning of the 60s of the 19th century**. Certain progress in the perfection of the working procedures was the implementation at the turn of the 19th and 20th centuries of the working procedure, **so-called, pegging** - the upper and lower parts of the shoes were connected using wooden pegs. [9]

From the technological viewpoint of show making it is necessary to mention that it has undergone substantially quick progress. The basis is that **up to the outbreak of the First World War shoes were mainly made by hand**, however after the war, i.e. after 1918, the use of **machines** in the shoe factories gradually started and subsequently increased. **The use of machinery in production replaced about 4 to 5 workers**. Production itself was divided into two parts, i.e. **production of the lower and upper part**, whereby **calf and cow hide** were mainly used for the upper part. The shoe bottoms were fitted with a pre-tanned lining.

In the shoe factory, **every worker had his specific function**, and for this reason, their contemporary markings survive, e.g. **slicers** (used to slice the upper parts), **splitters** (manual finishing of the bottom edges using the paring-knife), **gluers**, **grafters** (they did the sewing on the machines). The shoes then moved on to the "**ostráci**", who pegged or sewed the shoes. The last of the workers to receive the shoes were the **so-called polishers**, whose task was to polish the bottom. The tops of almost finished shoes were also polished in the factory and transferred to dispatch after meticulous inspection**. Crocodile leather** was also used to make **the most luxurious shoes**. Shoe manufacturing **was naturally seasonal**, and for this reason, the summer shoes were manufactured from January to June and the production of winter and hiking shoes was done up to the end of the year. [10]

Fig. 5. Shoe production at the turn of the 19th and 20th centuries. (The Municipal Museum of Skuteč)

About the author...

František Krejčí was born on 4 April 1991 in the Municipality of Krouna not far from the town of Skuteč, where he attended the Gymnasium of the Sovereign Military Order of Malta for four years. In the period 2010 to 2016, he studied at the Faculty of Education of UHK and accomplished both the bachelor's and subsequently the master's studies in History and the German Language with education major. From 2016 he is an intern doctoral student of the History Institute of the Philosophical Faculty of the UHK with specialization in the economic and social history of industrial enterprises in the modern Pardubice Region.

4. Resources

Bibliography:
[1] DOLEŽÁLEK, J. *Vlastivědný průvodce Skutčí a okolím.* Skuteč, 1967, p. 19.
[2] JETMAR, V. *Skutečské obuvnictví.* In: Vysokomýtsko. Vysoké Mýto, 1931, p. 393–394.
[3] FROLÍK, J. a kol. *Historie a současnost podnikání na Chrudimsku*. Žehušice, 2004, p. 122.
[4] FROLÍK, J. a kol. *Historie a současnost podnikání na Chrudimsku*. Žehušice, 2004, p. 123.
[5] JETMAR, V. *Skutečské obuvnictví.* In: Vysokomýtsko. Vysoké Mýto, 1931, p. 394.
[6] VORÁČEK, E. a kol. *Dějiny Skutče.* Skuteč – Hradec Králové, 2011, p. 205–208.
[9] PŘÍHODOVÁ, E. *Stručné dějiny oborů: Textil, Oděvnictví, Obuvnictví.* Praha, 2004, p. 24–26.
[10] JETMAR, V. *Skutečské obuvnictví.* In: Vysokomýtsko. Vysoké Mýto, 1931, p. 395.

Archival resources:
[7] SOA Zámrsk. Krajský soud Chrudim (1850–1949).
Spisy firemní. *Bratři Rybičkové. Výroba obuvi Skuteč.* karton č.: 780, inv. č.: 1269.
[8] SOA Zámrsk. Krajský soud Chrudim (1850–1949).
Spisy firemní. *Fr. Hošpes Skuteč.* karton č.: 781, inv. č.: 1269.

Amazonia and hydropower plants: Urban networks, decrease of the forest and spatial inequalities

Nádia BITAR

Historical Laboratory of (Electro)Technology, FEE, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic
bitarmat@fel.cvut.cz

Abstract. *The Brazilian Amazonia region, identified by its rain forest and home of 12% percent of the population of the country is focus of a rich scientific and intellectual production. For decades now, the matrix of the Brazilian geography and geopolitics has been building a solid interpretation of the relationship regarding the region, the national state, territory and development. This article aims to contemplate a few aspects that discuss localities and environmental and social impacts related to the emergence of a new urban network that has become strategic for the industry of hydroelectricity production. The case studied highlights issues that exposes how the political planning allied to the global energy market can influence the formation of unequal development in urban and agrarian space in the country. This process brings the "time of the urban space" the post-modernity and the global market language to places that weren't (in such an effective way) linked to those incomes. It can be verified, however, that such logic is repeated in other parts of the world. The purpose by now is to present a contribution supporting further conceptions for a methodological framework that highlights such issues in a global and deeper analyze.*

Keywords

Amazonia, Hydroelectric plants, urban network, energy industry, socio-environmental impacts, deforestation.

1. From geographic scale complexity to punctual inequalities

Currently, multinational companies operate large-scale works regarding the construction of hydroelectric plants and the implementation of transmission lines and substations. The following example under analyses is the Belo Monte Dam (the fourth world's largest electricity-generating hydropower plant) with the capacity of 11.233 MW (Norte Energia S/A 2019) and located in the central area of the Pará state, north of Brazil in the Amazonia forest. The construction of the dam, reservoir, power houses and further facilities started in 2011. The enterprise will house up 18 generating turbines (Francis type) and until nowadays the power plant isn't fully finished, only with 11 turbines operating.

Due to the building complex of this project, a new networks of towns wore formed in a specific zones chosen by the electric energy sector, the localities chosen by the enterprise as impacted of affected areas are Altamira (the main city that host the enterprise), Anapu, Brasil Novo, Senador José Porfírio and Vitória do Xingu, such towns are characterized in the environmental licenses process needed for the approval of the works, a concession of agreement by governmental institution that gives the environmental permits required to installation and operation of the hydropower plant. Today this places are found in critical situation once that they haven't been truly monitored and supported by the company. The Consortium Norte Energia S/A that is the owner of the business is definitely not complying with the legal steps to operate the hydroelectric plant.

The localities affected by the enterprise became an significant vector of deforestation and illegal activities towards the forests, they indirectly support and endure in urban networks which are built on these activities. In the same way, other towns and villages under the role of the power plant are changing really fast. In this reasoning, is possible to think about asymmetries of power once old socioenvironmental pacts are broken and other new orders are established in the territory. It is central for this analyze to think about the need of certain towns to harbor the transmission lines needed for complete the operation of the electrical system.

In such a way, the impacts of a hydroelectric plant of the size of Belo Monte enterprise exceed the regional scale and are in fact interstate. If we think about the national interconnected system of distribution of electricity, there

are consequences that can be felt throughout the national territory.

Another important item to be considered is the question of scale in terms of the Amazonia region. It should be taking into account that some municipalities have a territorial extension that corresponds to entire countries. The case of Belo Monte Dam is emblematic since the enterprise is located mostly in the municipality of Altamira, that has the territorial extension of 159 696 km². In this sense, the host city can competes with other states of the federation. The mayor of Altamira happens to be a match in relation to governors. However, they don't have the same administrative structure, as well finances and policy framework to do so. Thus, an unequal regime of forces is established regionally and with reflexes in national scale.

By national law is recognized the obligation to publicize the directives of companies in relation to the implementation of their operational structures. Socioenvironmental measures that provide compensation and programs to mitigate the damages caused are also required for the installation and operation permits. However, the resolutions become diffused, as well as the content of what is publicized remains unsatisfactory. Thus, social organization is prevented, delaying a possible reaction by the democratic control forces, communities organizations or even by governmental institutions as well the formal justice outcomes. In the same way, academic research is compromised.

In Brazil, the formula by which the development of the Amazonia region occurs is using the discourse of the advance of "civilization" towards the unexplored lands and shaping the image of the "savage people" of the forest. This process doesn't change much since centuries and is expressed by the exploitation of natural resources, mining and logging stocks. In this logic the large dams planned for the region plays a central role.

A quick search in the newspaper reports about Altamira, shows that the number of homicides grew in an incredible way as the enterprise was been established (Altamira a few years ago had reach the status of the most violet locality in the world). The fact that the population of the city has grown spontaneously in a very short time (because of the power installation) cannot in itself, explain the vertiginous increase of violence, the analysis of such imbalance needs to be done with care. These areas are still little tuned with the globalized markets, are found apart of process of normalization of information and from the mass consumption lifestyle; still constitute themselves as little orchestrated by the logic of an urban society.

This characteristics should be notice regarding its social organization, its particularities and determinations generated in other cultural backgrounds, which count with other temporalities and territorialities to be expressed (the case of the indigenous people is the biggest example). The installation of a hydroelectric plant in such isolated areas, ("isolated" cause we have as standard to think about reality the globalized world) end up breaking with social organization and stimulating spatial asymmetries.

This discussion that is open up is about how the hydroelectric power plant represents the advance of the "technization" of the territory, wile is also opening space for the increasingly intensive advance of the exploitation of resources in a generalized way. Altamira is a city that embodies this reasoning once is suffering the reflexes of this new technical structure.

2. The impact analyses and the urban network

Analyses guided by a critical evaluation about the content of the licensing environmental studies published by the companies, as well as reports of the media and the government involved in the project is one way to reach the real status about the consequences of the enterprise. Same time, the Geographical Information System - GIS can set satellites images making possible to visualize chronologically how new roads into the forest and new zones of illegal mining is under way simultaneously, due to the proximity of the power plant site and its necessary infrastructures.

In the location of the Xingu´s Big Band, downstream of where is located the dam it is set a great diversity of fluvial channels (more than 100 km of diverse canals) that in turn is formed given the geodiversity of the site. The whole set of geological and geomorphological characteristics of soils as well the hydrography itself, also reflects the great diversity of the river´s environment. This very particular area presents also a great diversity of endemic fauna and flora of the Xingu river basin, more than notable.

Several are the reports that put the region - where the dam is now changing completely the landscape - as one of the most diversified region in terms of biodiversity (icthyofauna as example) in the world. Side by side among others areas with also enormous value. All this is traduced in resources to be explored, and for that, the development of the urban network is vital. It is a need once it makes possible to penetrate into the forest but end up happening without any organization or planning. Actually the Xingu´s Big Bend is drying, this because the original course of the river was deviated to fulfill the reservoir and supply the hydroelectric power plant. Indigenous people are suffering with the lack of fish and the all life chain sported by the river, same for the riverside population unable to navigate in those waters as before.

The regard towards the emergence of a new urban network and centralities boosted by those big works of infrastructure, needs to evaluate the process of establishment of nodal points of a technological network. Which in turn, sustains much of the country's energy consumption.

The Brazilian geographic school has developed an explanation centered in concepts as concentration and dispersion, the need for the fluidity of merchandise and of information. As well, the idea of "luminous places" explains the current territorial formation of the country and its inequality and difficulty in establishing an organized direction and a centralized governance. (Santos, 1993 Apud Cataia, 2013). Instead, what can be observed is the expansion of technological nets of transport and information. The presence of networks in disagreement with environmental laws and governmental plans, at the several executive levels (municipalities/state/federation) become an important issue. These networks are set up under divergent interests (with often hidden actors) but that plays an important role in explaining the current political crisis in Brazil.

The analyses about the emergence of centralities driven by the energy business need to be considered in relation to others federal government strategic main plans; as the roads, the mining and waterways and ports projects and the agricultural zones of expansion. On the other hand, the decadence of undersupplied regions found outside of the role of the territorial modernization plan, needs also attention. Places that appear as “in disagreement” with the territorial planning policy became tension points capable of unbalancing the "whole". We see with the Professor Cataia, that the territory is modernized aiming density and fluidity, but in an intensively selective way, which creates new and dismiss old inequalities.

The Amazonia rain-forest is daily threatened by the agribusiness development based in large properties and that operates through the “cattle-soybean binomial". is a pressure originating in the center-west where soybean production is highly developed and consolidated and expands toward the north, that is, the Amazonia biome. In this process, indigenous lands (such as the Xingu National Park) or protected lands by law as national forests, are invaded. Often with indigenous populations victimized by attacks. First, it is set the economy of irregular roads supported by the theft of the more valuable wood, passing to the withdrawal of the rest of the vegetation and carrying out the implementation of grassing lands. In a second moment, it is set the advance with the soybean plantation pushing cattle ranching toward forested areas. Illegal mining of gold in alluvial beds of the mains rivers is also part of that logic of occupation.

According to the characteristics of this process, asymmetries stands out. Is possible to grasp the distinct distribution of services along this urban network that is formed, as well as the distinct application of taxes and tax refunds (which will then be available to local authorities). The evaluation of the impacts of the hydroelectric, discusses how the social inequalities result no longer from the problems only observed in local scale, but rather respond by global links and necessities. In this reasoning, the repercussions of systemic dysfunctions may be reflected in locations thousands of kilometers away.

What can be verified in several studies is that most of the need for new hydroelectric plants and the generation of electricity in Brazil is explained to destine this resource for mineral exploration and not for human consumption, as is usually believed. In this sense, it is known that Brazil is among the countries that support its economy in the production of commodities of this kind. And that this is a necessity not necessarily thought of by the Brazilian society, but instead, it established and explained by the world market players, designating the country that exerts this productive purpose and so to remain at this position. It is imperative that it needs to be problematized and studied.

For the Amazonian universe of analyze to consider the scale accredited to the planning polices and actions on the territory are indispensable. Old plans for the "integration of the Amazonia" became known by the general public during the military dictatorship (1964-1988). Beck then, the central idea of the geopolitics and planning for the region was the construction of large hydroelectric plants. The military projects for the integration of the region into the rest of the national territory were set with the help of a well know ideology that was largely widespread: “to occupy so do not deliver”.

Another central aspect of the military plans for the Amazonia region were to build roads as a true principle for the occupation, to connect the zones that could not be transposed since the extremely sinuous rivers and the waterfall and rapids areas that would prevent the navigable transposition, as well the close forest stop the advance into the forest. At that time as still is today, the main connection of the forest as a whole is with Belém capital of the state of Pará that has always been the main port to drain the production from Amazonia.

It is now known that federal highways served as real propulsors and inaugurated the main outbreaks of deforestation and illegal occupation on the contiguous and remote forests. The famous Transamazon Highway road (which wasn’t finished as was planned) was perhaps the greatest threat vector for the forest population as well for the preservation of its nature. It is as road that leaves the Brazilian coast to cross the entire forest until the more remote western frontiers. It seems to us that there was no coincidence that one of its stops was in the region of Altamira (closest to the Belo Monte hydroelectric plant), being an event that gave momentum to the city beck then. At that time, the military government justified its construction in this region under the slogan “lands without men for men without lands".

The construction of the Transamazon Highway openly announced the indigenous genocide. It is possible to say that the 1960s, 1970s and 1980s delimited a specific period fundamental for the current configuration of the social landscape of the Xingu basin region. In these three decades it was observed the consolidation of economic development projects and the advance of the urban society the ideal of the “national society” over the indigenous territories. These large industries as the dam industry and the mining industry, manifest the history of nationalization

of the State of Pará, in which the main economic interest by the national state was grounded over the mineral exploration that would support the basic industries (steel) and all the post-war demands of the international market in the second half of the 20th century. The construction of massive hydroelectric plants and roads in strategic areas planned beck in that time are still the main projects to be installed, with little modification in its original design. Currently in 2019 it's imperative to quote this time because once again Brazil has a political configuration of a military government given the new presidency largely represented by the military forces, and what that means to the needs and plans to Amazonia, was already seen by Brazilian society.

As a partial conclusion and future outlook, the text expresses how ineffective to the communities - to the full realization the citizenship or assurance of human rights and the rights of the original people - can be the act of adopting norms and technologies that respond to the need of fluidity and speed of the industry, designed by and for the global market.

It seems interesting to strass that the act of proposing signs or marks by giving names to the objects and to space; or to all that is appropriated (materially and symbolically) by society, is not limited to the condition that everyone knows such signs; this act has to be recognized as meaningful by society. The proposition of territorial borders expresses this idea. In this sense, proposing or imposing meanings for “objects" contained in the territory implies relationships of power. Currently, the national and international organizations that works with the subject of preservation of the Amazonia that are monitoring the region already knows that the landmarks of environmental protection areas and the indigenous nations territories are no longer considered legitimate. A new logic of occupation of space is formed and part of this new reality is drawn by the hydroelectric industry.

Acknowledgements

The research described in the paper was supervised by Prof. PhD Marecela Efmertová at CTU under my first semester at the PhD program. I would like to thank all the help and support I had especially of har once a new and challenging phase of life begins. I would also like to thank the support of the Committee for the History of Electricity and Energy of Électricité de France S.A.- EDF, which also encourage the subject of study under development for the thesis by supporting the continuity of the research work as well to attend academic extension activities.

References

[1] CATAIA, M. in *Território e macrossistema elétrico nacional. Concentração do consumo e dispersão da produção hidroelétrica*. In Anais do X Encontro Nacional de Pós-graduação em Geografia. Campinas, 2013b, p. 1-11.

[2] CATAIA, M. in *Uso do território e federação: novos agentes e novos lugares. Diálogos possíveis e participação política*. In Scripta Nova (Barcelona). 2010, p. 331 (16)

[3] GOLDSMITH, E. & HILDYARD N. in *The Social and Environmental Effects of Large Dams. Wadebridge Ecological Centre, Wadebridge, Cornwall;* 400p. 1984. online available in: http://www.edwardgoldsmith.org/1038/.

[4] NORTE ENERGIA In UHE de Belo Monte. Available in: http://norteenergiasa. com.br/site/wpcontent/uploads/2019

[5] SANTOS, S. M. S. B.; HERNANDEZ F. M. (Orgs.). in *Análise Crítica do Estudo de Impacto Ambiental do Aproveitamento Hidrelétrico de Belo Monte* (Painel de Especialistas). Belém, 2009.

About Author

Nádia BITAR has a Degree in Geography from the University of São Paulo (2012) with Individual Graduation Work -TGI with emphasis on Human Geography. Has professional experience in national and multinational companies in the area of environmental analyses involving licensing and strategic studies specifically related to the evaluation of socio-environmental impacts. Has also worked with territorial and urban planning. Holds the diploma Master Erasmus Mundus Techniques, Heritage, Territories of Industry (2018) coordinated by the Université Paris 1 Panthéon-Sorbonne (France), and carried out in consortium with Università degli Studi di Padova - UNIPD (Italy), and the University of Évora - EU (Portugal), with mobility internship performed in the Faculty of Electrical Engineering in the Czech Technical University of Prague (CTU). During the two years of master's degree, was supported with the annual grant by the Comité d'Histoire de l'Electricité et de l`Energie d`EDF (Electricité de France) and concluded the dissertation regarding the subject of landscapes and hydropower plants. Currently work as an independent consultant in the areas of geography; heritage protection, landscape, history, territorial planning and evaluation of social and environmental impacts and is doing a PhD in the Historical Laboratory of (Electro)Technology, at the Faculty of Electrical Engineering in the Czech Technical University of Prague CTU, for which he also won the grant of the EDF´s comitee.

Jan Henniger of Eberk, Lawyer at the Services of the Prague Polytechnic

Mgr. Vít ŠMERHA
Doctoral Study Program History of technology at the Czech Technical University in Prague
Archives of the Czech Technical University in Prague, Zikova 2, 160 00 Prague, Czech Republic
Vit.Smerha@cvut.cz

Abstract. *Jan Henniger of Eberk was one of the members of the Provincial Committee and an Officer in Charge of the Prague Polytechnic's affairs. On April 9, 1832, he was appointed the second Director of the Polytechnic Institute of the Kingdom of Bohemia, i.e. on the same day on which his predecessor František Josef Gerstner, the first Director and a professor of mechanics, retired. His main task was to consolidate the management of the technical school and to enforce some major reform proposals to reorganize the curricula drafted already by his predecessor Gerstner. The most important achievement was the establishment of the first Real School in Prague under the auspices of the Polytechnic Institute, which served as a preparatory educational facility for the students intending to enter the Polytechnic. An equally important project was the establishment of new professorships for the subjects comprising natural history, agriculture and commodity expertise. A graduated lawyer himself, he significantly contributed to the remarkably smooth functioning of the school administration, and, thanks to him, such outstanding experts as e.g. the distinguished physicist and mathematician Christian Doppler, taught here.*

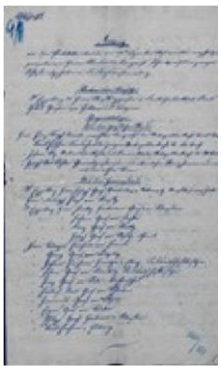
Keywords
Technical Education, Officer/Administration, Czech State Diet and Provincial Committee, Study Committee, Mathematics, Reforms, Study Plan.

1. Jan Henniger, an officer and lawyer, and his appointment as Director of the Polytechnic

Jan Henniger, the Freiherr (free lord) of Eberk, descended from a noble family. He was born in 1777 in Svolgim. He studied law at Prague University and since 1802 he worked in the civil service of the Provincial Governorate. Prior to joining the Polytechnic, he had long held the position of a Provincial Committee Associate Justice, starting from 1805 onwards, so he had a solid overview of the administrative agenda generated in Gerstner's time. He himself worked as an Officer in Charge of the Polytechnic's Affairs.

The election and appointment of Jan Henniger as Director of the Polytechnic was discussed by the Czech State Diet during the negotiations on the reorganization of the school. As early as on December 17, 1827, the State Diet decided that, after the retirement of František Josef Gerstner, the post of the Director of the School would be filled by an associate member of the Provincial Committee. Jan Henniger was chosen from all the members of the Committee as he was best aware of the school's internal workings. His appointment as Director of the Polytechnic in Prague was perceived as being just a temporary one, which was best evidenced by the Interim Director title he received.

Nonetheless, he managed the Prague Polytechnic for almost 13 years to be only replaced by František Chanovský in 1845. Jan Henniger certainly could not stand up to or match the professional, pedagogical and organizational qualities of his predecessor, F. J. Gerstner, yet his specific contribution consisted in reintroducing the order within the administrative processes after Gerstner's departure, when the Directorate was overwhelmed with documents related to the completion of important, unimplemented yet Gerstner's school reorganization proposals. Jan Henniger possessed some important prerequisites to hold the Director post. As an Associate Member of the Provincial Committee and an Officer in Charge of the Polytechnic's Affairs, he had gained a wealth of invaluable experience. He had also performed as the Chief Executive Officer of a newly founded Česká spořitelna (Böhmische Sparkasse). He was a founding member of the Union for the Encouragement of Industry in Bohemia and a member of the Patriotic-Economic Society. He set up one of the first sugar mills at his estate of Rtišovice, Smolotely and Bohostice in Příbram County.



2. School consolidation under Henniger's directorship and the enforcement of Gerstner's reforms

His appointment was associated primarily with the advancement of Gestner's key reforms. The curricula reorganization, which Gerstner had been planning on, could not take place due to his protracted disputes with the Study Committee. The Committee did not approve of the reorganization until May 28, 1832, after Gerstner's retirement and a month before his death.

In this light, the key project for the Polytechnic, successfully promoted by Henniger, was the establishment of regular professorships for the newly established branches: agriculture, natural history and commodity expertise. The professorship auditioning procedures were most challenging and took almost six years to complete. The Director eventually successfully vindicated 6 regular professorships at the Court Study Committee. One of the first implemented changes concerned the professorship of mechanics, which, after F.

J. Gerstner's departure, was being deputised by Karel Wersin. The mechanics curriculum was extended by adding lectures in physics. Elementary mathematics and practical geometry underwent significant developments, too. The essentials of mathematics were of crucial importance for the technical subjects' teaching. In this respect, Director Henniger took a fundamental stance in defence of mathematics and rejected the proposal tabled by the Court Study Committee in 1832 to remove mathematics from the polytechnic curricula and to only preserve the elementary mathematics subject at the Real School. Henniger not only asserted the mathematics teaching should be preserved, but even established an independent elementary mathematics and practical geometry professorship in 1839. It was this very subject that was later taught, first at the Real School and later on at the Polytechnic, by Christian Doppler, an Austrian physicist and mathematician whom Henniger had invited here. Doppler also lectured on higher mathematics at the Polytechnic between the years 1836 and 1837. However, he did not carry on with the lectures because the Study Committee cancelled them on the grounds of their parallel existence at the University.

Civil engineering was taught by Karl Wiesenfeld. He had been deputising here for almost nine years before, to be finally appointed regular professor of civil engineering in 1838, having passed a lengthy selection procedure. Karel Balling was undoubtedly an important figure at the Polytechnic. As

early as aged 19, he was chosen by Director Gerstner to become a deputising chemistry professor. In 1835, he was appointed regular professor of chemistry. The newly established independent branches of agriculture, natural history and commodity expertise were taught by Josef Lumbe, who had been appointed professor already in 1832 by František Gerstner, and by mineralogist and geologist František Zippe, who became professor of natural history and commodity expertise in 1835.

One of the proposed Gerstner's reforms, which had been tabled already in 1810, was the establishment of the first Real School in Prague as part of the Polytechnic. The proposal was passed by the Emperor no sooner than in 1832 and the actual teaching started in the school year 1833/34. This move brought about a need for new construction activity to arrange for adequate Real School classroom spaces. The most radical move in the reorganization of the Polytechnic thus was that of establishing additional Real Schools in Liberec and Rakovník. The plan was that theoretical and practical training for the future Polytechnic students would be taking place here. The Polytechnic was particular about the professional standards of the newly established secondary Real Schools. An important prerequisite was the vocational internships of the secondary school teachers who had to pass demanding auditions under the supervision of the lecturing Polytechnic professors, who were also members of the selection boards. It was the future teachers' duty to corroborate their previous work histories and educational backgrounds.

3. Henniger's collaborators and patrons of the Polytechnic

Henniger himself was a bureaucrat in the true sense of the word, and he demanded the same approach to dealing with the official agenda from collaborators. He shared his Director's responsibilities with his subordinates, too. He delegated specific tasks to certain members of his office and to the professor staff, e.g. Karel Balling was in charge of the bookkeeping activities and looked after the library, while Josef Lumbe and František Zippe supervised the collections, material equipment and maintenance and operations of the buildings. As far as the vocational part of teaching was concerned, Henniger had new seminars and laboratories established to bring the theoretical lecturing closer to the technical practice, as well as he required that the teachers, being experts representing the Polytechnic Institute, should draw up expert reviews and testimonials for industrial enterprises, thus bringing theory closer to practice and preventing teachers from being encapsulated in their academic school environments. This sense of modernity, which we are requiring from today's technical education, can be viewed as a substantial change in the life of the technical community and the whole society of that time. Henniger also tried to avert the extracurricular activities of the Polytechnic's pedagogues. He urged them to quit any extra jobs they may have been having, to devote fully to student teaching and to refrain from performing any other non-professional or non-teaching activities outside the Polytechnic within any commensurable extents. It was at his suggestion that all professors' paid workloads were increased. An exception was František X. M. Zippe, professor of natural history and commodity expertise, who performed unpaid roles as a Custodian in the Patriotic Museum and a Secretary of the Patriotic-Economic Society. Director Henniger launched an interesting initiative aiming at improving the studies at the Polytechnic by presenting a new curriculum concept. Its aim was to to split the study program into two directions: Mathematics and Natural Science. Mathematics was to comprise lectures in physics, elementary mathematics and practical geometry. These would also cover mechanical engineering, hydraulics and civil engineering. In addition to natural science as such, the Natural Science line was to include mineralogy, commodity expertise, agriculture, chemistry, and property management as well. However, this study system failed to make the students abide by its

curriculum. The project remained at the trial stage, with its actual implementation taking place no sooner than when the so-called Organic Statute was published in 1863. Another innovative act was the introduction of tuition fees in 1836. This measure was originally intended as a financial aid facilitating the reorganisation of the studies. However, it had a positive effect in enhancing the students' activity towards passing their examinations, as a prerequisite had been laid down of having to pass the examinations to qualify for exemption from the tuition fees.

4. Conclusion.

Director Henniger unveiled his intentions in a speech delivered on 10 March, 1835, making references to František J. Gerstner, his predecessor in the Directorial position. During his speech, he introduced the aims and priorities of the school in general, with an emphasis being put on educating professionals in the agricultural and forestry economy rather than in the technical fields. He was awarded the Imperial Knight Order of Leopold I in 1836 and had been named Austro-Hungarian Imperial Chamberlain even earlier.

Director Henniger achieved his most remarkable success when, through an Imperial Decree of October 24, 1843, the Polytechnic in Prague was granted the same rights as the Polytechnic in Vienna. His failures included the inability to complete the reformative efforts started already under his predecessor, F. J. Gerstner. The most significant contribution to the reorganization of the school thus remained the consolidation of the school's administrative processes and the establishment of a fundamental educational structure, including manning all professorships as a standard practice and engaging prominent experts..

Henniger was also involved in charitable activities, he was a contributor to and a member of the St. John the Baptist Orphanage Committee in Lesser Town, Prague. When Henniger's performance as Associate of the Provincial Committee ceased in 1842, he remained in the Director's office for three more years. Then, in the course of the study year 1844/45, he decided to resign, retire and pursue farming at his estate.

Acknowledgements

The research described in this paper has been supervised by Prof. M. Efmertová, Supervisor at FEE CTU in Prague, and supported by the Student Grant Agency.

References

[1] Archives of the CTU, Funds: Polytechnic Institute, 1827, 1832, k. 15, 20.

[2] Karl Jelinek, The gentry - Polytechnic Institute to Prague, Prague 1856, pp. 199-201; Franz Stark, The German Technical University in Prague, 1806-1906, Prague 1906, 8-10.

Albert Vojtěch Velflík, The history of the Technical Teaching in Prague, I/1, Praha 1906, s. 251 - 252

[3] František Jílek – Václav Lomič, The history of the CzechTechnical University, I/1, 1973, Prague s. 288 – 302;

[4] Magdaléna Tayerlová, The period of administrative management and consolidation of the Polytechnic Institute, in: Czech Technical University, Prague 2004, s. 48 – 51..

Vít Šmerha, On the Paper of the Prague Polytechnic School around 1832, in: Zdeněk Hojda – Hana Pátková, Pragmatic documents in the context of legal and administrative, Faculty of Art of the Charles University, Prague 2008, s. 275 – 287.

Mgr. Vít ŠMERHA was born in Prague. He studied at the Faculty of Philosophy at Charles University in Prague. He works at the Archive of the CTU Prague.

The Masaryk Academy of Labour and its Contribution to the Development of Technically Skilled Workers

Vít HOLEČEK

Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 166 27 Praha, Czech Republic

vitholecek@seznam.cz

Abstract. *The Masaryk Academy of Labour (MAL) was founded in 1920 as the first academic institution for Czech (Czechoslovakian) technicians and engineers. It was the centre of technical work whose aim was to organise it so that the skills of all people would be used as well as the natural resources of Czechoslovakia in an economical way. Due to the change of political system, the institution became part of the Czechoslovakian Academy of Science in 1952. As an academic technical institution, the Masaryk Academy of Labour influenced also the development and adjustment of high school education and university technical education in the form of independent proposals. Thanks to its Commission for the Democratisation of Education, it supported the professional development of technically skilled workers in the form of practical handbooks and lectures. There was also the Commission for the Problem of Unemployed Technically Skilled Workers, whose aim was to find how to use the abilities of unemployed technicians. Last but not least, it also financially and administratively supported work placement schemes and academic residency programmes of the Czech engineers in American major industrial plants, such as Ford's Plant in Detroit, etc. By all these activities, the Masaryk Academy of Labour contributed to the development of the engineering elite in the newly formed Czechoslovakian Republic. The aim of the article will be to introduce the activities of the above mentioned Commission for the Democratisation of Education.*

Keywords

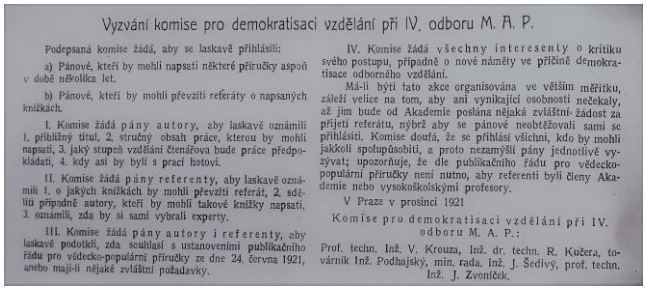
Masaryk Academy of Labour, Technical Institute, History of Science and Technology, Engineers, Education, Technical Workers.

1. Supplementary Education for Technically Skilled Workers

Fig. 1 The seat of the Masaryk Academy of Labour in the original residency of the Kinsky noble family in the Old Town Square, Prague

In the winter of 1918, Ing. Dr. techn. Rudolf Kučera lectured on the themes "The Social Question and the Third Condition of Life" and "The Possibilities of Democratic Equality" for the Association of Engineers and Architects. [1] Already in that time, these questions were attractive for the Preparatory Committee of the MAL and, after that, they also appeared in the form of the draft bill on the foundation of the academy, i.e. to use the skills of all people as well as the natural resources of the Czechoslovakian Republic. In 1920, this act was adopted by all parties in the parliament. That is why the MAL was not focused only on the development of engineering but also on the improvement of social conditions of workmen.

In November 1921, the 4th Class and the 4th Division of Mechanical and Electrical Engineering issued the declaration asking Czechoslovakian engineers and professors to publish technical handbooks for workers. In this way, the 4th division of the MAL wanted to spread technical education to all classes of the nation, in other words to "democratise technical education" (Fig. 2 and 3). [1]



Vyzvání komise pro demokratisaci vzdělání při IV. odboru M. A. P.

Podepsaná komise žádá, aby se laskavě přihlásili:

a) Pánové, kteří by mohli napsati některé příručky aspoň v době několika let.

b) Pánové, kteří by mohli převzíti referáty o napsaných knížkách.

I. Komise žádá pány autory, aby laskavě oznámili 1. přibližný titul, 2. stručný obsah práce, kterou by mohli napsati, 3. jaký stupeň vzdělání čtenářova bude práce předpokládati, 4. kdy asi by byli s prací hotovi.

II. Komise žádá pány referenty, aby laskavě oznámili 1. o jakých knížkách by mohli převzíti referát, 2. sdělili případně autory, kteří by mohli takové knížky napsati, 3. oznámili, zda by si sami vybrali experty.

III. Komise žádá pány autory i referenty, aby laskavě podotkli, zda souhlasí s ustanoveními publikačního řádu pro vědecko-populární příručky ze dne 24. června 1921, anebo mají-li nějaké zvláštní požadavky.

IV. Komise žádá všechny interesenty o kritiku svého postupu, případně o nové náměty ve příčině demokratisace odborného vzdělání.

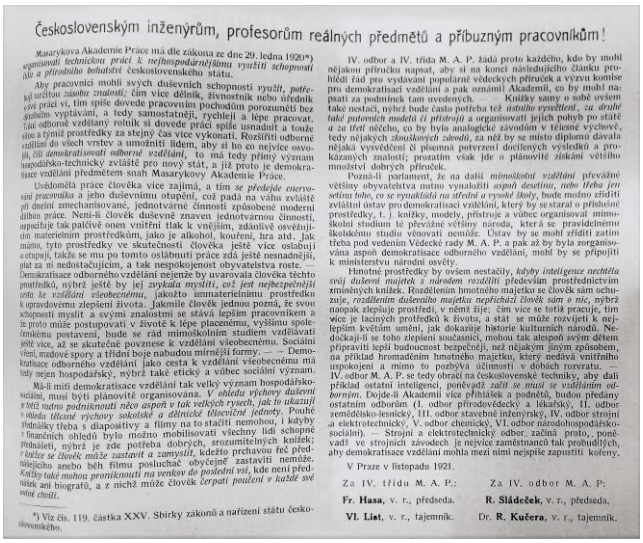
Má-li býti tato akce organisována ve větším měřítku, záleží velice na tom, aby ani vynikající osobnosti nečekaly, až jim bude od Akademie poslána nějaká zvláštní žádost za přijetí referátu, nýbrž aby se pánové neobtěžovali sami se přihlásiti. Komise doufá, že se přihlásí všichni, kdo by mohli jakkoli spolupůsobiti, a proto nezamýšlí pány jednotlivě vyzývat; upozorňuje, že dle publikačního řádu pro vědecko-popularní příručky není nutno, aby referenti byli členy Akademie nebo vysokoškolskými profesory.

V Praze v prosinci 1921

Komise pro demokratisaci vzdělání při IV. odboru M. A. P.:

Prof. techn. Inž. V. Krouza, Inž. dr. techn. R. Kučera, továrník Inž. Podhajský, min. rada. Inž. J. Šedivý, prof. techn. Inž. J. Zvoníček.

Fig. 2 The appeal of the MAL to academic public to publish specific handbooks for workers [4].



Československým inženýrům, profesorům reálných předmětů a příbuzným pracovníkům!

Masarykova Akademie Práce má dle zákona ze dne 29. ledna 1920*) *organisovati technickou práci k nejhospodárnějšímu využití schopností a přírodního bohatství československého státu.*

Aby pracovníci mohli svých duševních schopností *využíti, potřebují určitou zásobu znalostí;* čím více dělník, živnostník nebo úředník své práci ví, tím spíše dovede pracovním pochodům porozuměti bez zbytného vyptávání, a tedy samostatněji, rychleji a lépe pracovat. Také odborné vzdělaný rolník si dovede práci spíše usnadnit a touže sílou a týmiž prostředky za stejný čas více vykonati. Rozšíření odborné vzdělání do všech vrstev a umožnití lidem, aby si ho co nejvíce osvojili, *čili demokratisovati odborné vzdělání,* to má tedy přímý význam hospodářsko-technický zvláště pro nový stát, a již proto je demokratisace vzdělání předmětem snah Masarykovy Akademie Práce.

Uvědomělá práce člověka více zajímá, a tím se *předejde enervování pracovníka* a jeho duševnímu otupění, což padá na váhu zvláště při dnešní zmechanisované, jednotvárné činnosti způsobené moderní dělbou práce. Není-li člověk duševně znaven jednotvárnou činností, nepociťuje tak palčivě onen vnitřní tlak k vnějším, zdánlivě osvěžujícím materiálním prostředkům, jako je alkohol, kouření, hra atd. Jak známo, tyto prostředky ve skutečnosti člověka ještě více oslabují a otupují, takže se mu po tomto oslabnutí práce zdá ještě nesnadnější, plat za ni nedostačujícím, a tak nespokojenost obyvatelstva roste. — Demokratisace odborného vzdělání nejenže by uvarovala člověka těchto prostředků, nýbrž ještě by jej *zvykala myšlití, což jest nejbezpečnější cesta ke vzdělání všeobecnému,* jakožto immaterielnímu prostředku k opravdovému zlepšení života. Jakmile člověk jednou pozná, že svou schopností myslit a svými znalostmi se stává lepším pracovníkem a že proto může postupovati v životě k lépe placenému, vyššímu společenskému postavení, bude se rád mimoškolním studiem vzdělávati ještě více, až se skutečně povznese k vzdělání všeobecnému. Sociální tření, mzdové spory a třídní boje nabudou mírnější formy. — Demokratisace odborného vzdělání jako cesta k vzdělání všeobecnému má tedy nejen hospodářský, nýbrž také etický a vůbec sociální význam.

Má-li míti demokratisace vzdělání tak velký význam hospodářsko-sociální, musí býti plánovitě organisována. *V ohledu výchovy duševní* a vůli *mutno podnikouti něco aspoň v tak velkých rysech, jak to ukazují* v oblasti *tělesné výchovy sokolské a dělnické tělocvičné jednoty.* Pouhé přednášky třeba s diapositivy a filmy na to stačiti nemohou, i kdyby po finančních ohledů bylo možno mobilisovati všechny lidi schopné přednášeti, nýbrž je zde potřeba dobrých, srozumitelných knížek. *Knihu se člověk může zastavit a zamyslit,* kdežto prchavou řeč přednášejícího anebo běh filmu posluchač obyčejně zastaviti nemůže. *Knížky také mohou proniknouti na venkov do posledních vsí,* kde není přednášek ani biografů, a z nichž může člověk čerpati poučení v každé své volné chvíli.

*) Viz čís. 119. částka XXV. Sbírky zákonů a nařízení státu československého.

IV. odbor a IV. třída M. A. P. žádá proto každého, kdo by mohl nějakou příručku napsat, aby si na konci následujícího článku prohlédl řád pro vydávání populárně vědeckých příruček a výzvu komise pro demokratisaci vzdělání a pak oznámil Akademii, co by mohl napsati za podmínek tam uvedených. — Knížky samy o sobě ovšem také nestačí, nýbrž bude často potřeba *též ústního vysvětlení, za druhé také putovních modelů či přístrojů* a organisovati jejich pohyb po státě *a za třetí* něčeho, co by bylo analogické závodům v tělesné výchově, tedy nějakých *zkouškových závodů,* za něž by se místo diplomů dávala nějaká vysvědčení či písemná potvrzení dočtených výsledků a prokázaných znalostí; prozatím však jde o plánovité získání většího množství dobrých příruček.

Pozná-li parlament, že na další *mimoškolní vzdělání* převážné většiny obyvatelstva nutno vynaložiti *aspoň desetinu, nebo třeba jen setinu toho, co se vynakládá na střední a vyšší školy,* bude možno zříditi zvláštní ústav pro demokratisaci vzdělání, který by se staral o příslušné prostředky, t. j. knížky, modely, přístroje a vůbec organisovat mimoškolní studium té převážné většiny národa, která se pravidelnému školskému studiu věnovati nemůže. Ústav by se mohl zříditi zatím třeba pod vedením Vědecké rady M. A. P. a pak až by byla zorganisována aspoň demokratisace odborného vzdělání, mohl by se připojiti k ministerstvu národní osvěty.

Hmotné prostředky by ovšem nestačily, *kdyby inteligence nechtěla svůj duševní majetek k národnímu rozdílit* především prostřednictvím zmíněných knížek. Rozdělením hmotného majetku se člověk sám ochuzuje, *rozdělením duševního majetku nepřichází člověk sám o nic,* nýbrž naopak zlepšuje prostředí, v němž žije: čím více se totiž pracuje, tím více je lacinějších prostředků k životu, a stát se může rozvíjeti k nejlepším květům umění, jak dokazuje historie kulturních národů. Nedokaží-li se toho zlepšení současníci, mohou tak alespoň svým dětem připravit lepší budoucnost bezpečnější, než nějakým jiným způsobem, na příklad hromaděním hmotného majetku, který nedává vnitřního uspokojení a mimo to pozbývá účinnosti v dobách rozvratu. — IV. odbor M. A. P. se tedy obrací na československé techniky, aby dali příklad ostatní inteligenci, poněvadž *začít se musí se vzděláním odborným.* Dojde-li Akademii více přihlášek a podnětů, budou předány ostatním odborům (I. odbor přírodovědecký a lékařský, II. odbor zemědělsko-lesnický, III. odbor stavebně inženýrský, IV. odbor strojní a elektrotechnický, V. odbor chemický, VI. odbor národohospodářsko-sociální). — Strojní a elektrotechnický odbor začíná proto, poněvadž ve strojních závodech je nejvíce zaměstnanců tak probudilých, aby demokratisace vzdělání mohla mezi nimi nejspíše zapustiti kořeny.

V Praze v listopadu 1921.

Za IV. třídu M. A. P.: **Fr. Hasa**, v. r., předseda. **Vl. List**, v. r., tajemník.

Za IV. odbor M. A. P: **R. Sládeček**, v. r., předseda. Dr. **R. Kučera**, v. r., tajemník.

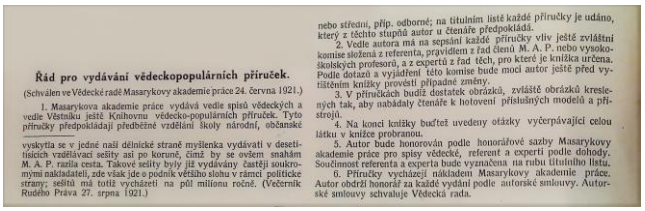
Fig. 3 The announcement of the MAL for Czechoslovakian academic public about the intention to educate technically skilled workers [4].

In order to solve these questions, the Commission of the Democratisation of Education was established. [1], [2] The following experts were elected to be its members: prof. Ing. Dr. techn. Václav Krouza, Ing. Dr. techn. Rudolf Kučera, the manufacturer Ing. Otakar Podhajský (Fig.4), minister counsellor Ing. J. Šedivý and prof. Ing. Dr. techn. Jan Zvoníček (Fig. 5).

Fig. 4 Ing. Otakar Podhajský (1876-1940), the engineer and manufacturer from Hostivař, Prague [5], [6].

Fig. 5 Prof. Ing. Dr. techn. Jan Zvoníček (1865-1926), the rector of the Czech Technical University in Brno and Prague. He constructed the so called Zvoníček's gas-steam turbine. He belongs to the founders of Czechoslovakian engineering. [7]., [8].

The aim of the democratisation of education was to open the way for raising the level of education and thus also the social rank for technicians of all social classes. According to the experts from the 4th division of the MAL, the mechanical work could cause mental dulling. "Such workers do not enjoy their work and, because of that, they are addicted to alcohol, smoking, and hazard." [1] The Commission wanted to make the workers more interested in their work, enhance their capacity to think, and improve their living conditions. With this purpose, the MAL issued regulations on Technical Handbooks Publication (Fig. 6).



Řád pro vydávání vědeckopopulárních příruček.

(Schválen ve Vědecké radě Masarykovy akademie práce 24. června 1921.)

1. Masarykova akademie práce vydává vedle spisů vědeckých a vedle Věstníku ještě Knihovnu vědecko-popularních příruček. Tyto příručky předpokládají předběžné vzdělání školy národní, občanské vyskytla se v jedné naší dělnické straně myšlenka vydávati v desetitisících vzdělávací sešity asi po koruně, čímž by se ovšem snažám M. A. P. rašila cesta. Takové sešity byly již vydávány častěji soukromými nakladateli, zde však jde o podnik většího slohu v rámci politické strany; sešitů má totiž vycházeti na půl miliónu ročně. (Večerník Ruského Práva 27. srpna 1921.)

nebo střední, příp. odborné; na titulním listě každé příručky je udáno, který z těchto stupňů autor u čtenáře předpokládá.

2. Vedle autora má na sepsání každé příručky vliv ještě zvláštní komise složená z referenta, pravidlem z řad členů M. A. P. nebo vysokoškolských profesorů, a z expertů z řad těch, pro které je knížka určena. Podle dotazů a vyjádření této komise bude moci autor ještě před vytištěním knížky provésti případné změny.

3. V příručkách budiž dostatek obrázků, zvláště obrázků kreslených tak, aby nabádaly čtenáře k hotovení příslušných modelů a přístrojů.

4. Na konci knížky buďtež uvedeny otázky vyčerpávající celou látku v knize probranou.

5. Autor bude honorován podle honorářové sazby Masarykovy akademie práce pro spisy vědecké, referent a expertí podle dohody. Součinnost referenta a experta bude vyznačena na rubu titulního listu.

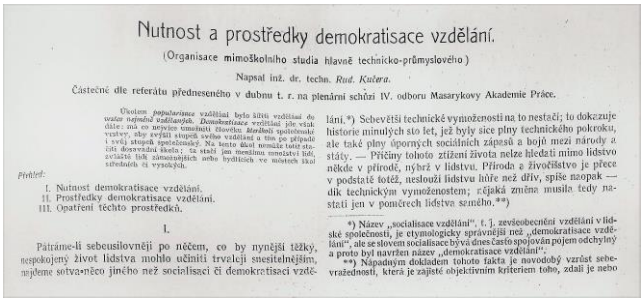
6. Příručky vycházejí nákladem Masarykovy akademie práce. Autor obdrží honorář za každé vydání podle autorské smlouvy. Autorské smlouvy schvaluje Vědecká rada.

Fig. 6 Regulations on Technical Handbooks Publication [4].

The 4th division of the MAL was the first division responsible for the democratisation of education as it was supposed that, in machinery plants, there was the highest number of workers who were in need of mental enhancement. [1]

The aim of the MAL was to offer technical education to everybody regardless their property, age, and occupation. The study was private without any need to study high school, vocational school, or other similar type of school. The experts of the MAL point out: "Those who started to work in young age after secondary school are invited to private study in order to advance to the higher working class." [1] The cooperation was established with the Ministry of Education and National Enlightenment, Ministry of Public Works, Ministry of Railways, universities and colleges, Engineering Chamber, Association of Engineers and Architects, Awareness Raising Association, Labour Academy, Central Labour School, etc. The professors from high schools and universities then tested the gained knowledge in order to make the workmen competitive and motivated to achieve

better results in their subjects. The first request for publication published in newspapers and academic journals was answered by thirty respondents. This theme was also the subject of a separate booklet called "The Necessity and Means of the Democratisation of Education" [4] written by Ing. Dr. techn. Rudolf Kučera and published as a special issue of Technický obzor – the journal of Czechoslovakian engineers and architects (Fig. 7).



Nutnost a prostředky demokratisace vzdělání.

(Organisace mimoškolního studia hlavně technicko-průmyslového)

Napsal inž. dr. techn. *Rud. Kučera.*

Částečné dle referátu přednesného v dubnu t. r. na plenární schůzi IV. odboru Masarykovy Akademie Práce.

Úkolem popularisace vzdělání bylo šířiti vzdělání do vrstev nejvíce vzdálených. Demokratisace vzdělání jde však dále: má co nejvíce umožniti člověku středních společenských vrstev, aby zvýšil stupeň svého vzdělání a tím po případě i svůj stupeň společenský. Na tento úkol nemůže totiž stačiti dosavadní škola; ta stačí jen menšímu množství lidí, zvláště lidí zámožnějších nebo bydlících ve městech nad středními či vyššími.

Přehled:

I. Nutnost demokratisace vzdělání.
II. Prostředky demokratisace vzdělání.
III. Opatření těchto prostředků.

I.

Pátráme-li sebeusilovněji po něčem, co by nynější těžký, nespokojený život lidstva mohlo učiniti trvalejí snesitelnějším, nejdeme sotva-něco jiného než socialisaci či demokratisaci vzdělání.*) Schevětší technické vymožeností na to nestačí; to dokazuje historie minulých sto let, jež byly sice plny technického pokroku, ale také plny úporných sociálních zápasů a bojů mezi národy a státy. — Příčiny tohoto ztížení života nelze hledati mimo lidstvo v přírodě, nýbrž v lidstvu. Příroda a živočišstvo je přece někde v přírodě, nýbrž v lidstvu. Příroda a živočišstvo je přece v podstatě totéž, neslouží lidstvu hůře než dřív, spíše naopak — dík technickým vymožeností; nějaká změna musila tedy nastati jen v poměrech lidstva samého.**)

*) Název „socialisace vzdělání", t. j. zevšeobecnění vzdělání v lidské společnosti, je etymologicky správnější než „demokratisace vzdělání", ale se slovem socialisace bývá dnes často spojován pojem odchylný a proto byl navržen název „demokratisace vzdělání".

**) Nápadným dokladem tohoto fakta je novodobý vzrůst sebevražednosti, která je zajisté objektivním kriteriem toho, zdali je nebo

Fig. 7 The article "The Necessity and Means of the Democratisation of Education" by Ing. Dr. techn. Rudolf Kučera [4].

It was supposed that young people wanted to start with practical life rather than with studies and that the social order would change. The idea was that the young people were to work manually first and only then, during their life, they would aim at higher working positions requiring higher education. It was desirable that school and life were connected, which followed the ideas of Jan Amos Comenius (1592-1670), the teacher of nations. The struggle of the MAL to democratise education was based on the words of President T. G. Masaryk (1850-1937) from the time when the MAL was founded. He remarked that "this academy is in line with the pansophic aspirations of Comenius that all work in the state needs to be organised on scientific bases and that **there will be no economic socialisation without the socialisation of education**". [1] Already before the World War I, the daily newspaper of the social democratic party Právo lidu wrote in its editorial that "the contemporary student does not study to know something but to gain a privileged position in the state". [1]

One of the practical examples of the activities of the Commission for the Democratisation of Education was the counting course, which took part under the auspices of the 4th division of the MAL from November 23, 1929 to March 1, 1921 upon the request of the Educational Division of the Union of Lathe Operators. The lectures took place once a week in the lecture room of the Department of Mechanical Technology at the College of Mechanical and Electrical Engineering, CTU. The course taught about "the shapes and systems of most frequent threads and the choice of interchangeable gears for thread cutting of all kinds and in all cases of lathing". [1] The emphasis was also put on Czech mechanical terminology, which Czech workers of the First Republic were not so much familiar with. Based on this lecture, there was also the first popular scientific publication of prof. Ing. RTDr. Otakar Grössl (1881-1941) - (*the professor of mechanical technology and tooling construction at CTU, Prague and the founder of the Department of Tooling Studies at CTU, Prague).* At the CTU, there was also a course of thermal management, and one part of the programme was conducted by the Coal Institute, which was organised by the MAL.

In October 1921, the 4th division of the MAL submitted the plan concerning the democratisation of the academic education to President T. G. Masaryk. At the meeting of the Commission for the Democratisation of Education, one of its members Ing. R. Žižka proposed to establish an evening school at the MAL called "Masaryk's School of Work", "Masaryk's School for Workers" or "Comenius' School of Work", whose aim would have been to educate technical employees without academic education aged from 21 to 40. The study should have been divided into 3 or 4 semesters, and it should have offered also general education, which the workers could not gain in their young age. There should have been also a department for women, who would have learnt about economical housekeeping. [1]

In that time, the similar effort was made by the City et Guilds of London Institute. The experts in the Commission for the Democratisation of Education of the MAL were also thinking of introducing correspondence teaching like it was at the International Correspondence School in Scranton, Pennsylvania, USA. This school was active in various parts of Europe, in London, Paris or Madrid. The owner of this International Correspondence School was the International Textbook Company, which cooperated with the International Educational Publishing Company. The International Correspondence School in Scranton was founded in 1891. Its first aim was the education of miners, who wanted to gain the necessary qualification for the position of mine inspector or mine foreman as required by the state of Pennsylvania. The success of the mining course was the inspiration to organise more courses like engineering drawing and electrical engineering; the number of students was growing fast. In 1920s, the number of the core courses devoted to engineering had risen up to 200. In 1918, the school had more than 2 millions of students. [1]

Besides the publication of the rules for academic texts, the MAL had also publication rules for popular scientific texts. The publication rules for popular scientific booklets of the MAL established that any author fulfilling the conditions of these rules could be the author of such text. The officers from the MAL or university professors were there to cooperate with each author and control the text from the factual point of view. A co-author could be also a workman, who asked questions to demonstrate if the content of the publication is clearly explained. It was significant that the publications contained a lot of pictures, which could guide the reader to construct the respective models and devices. The MAL wanted also to introduce new technological innovations by devices, models, films, or slides. There was also a plan to create a separate Institution for the Democratisation of Education, which would have published booklets, cared about models, devices, and their movement in the state. It would have

organised educational tours with films and respective courses, and it would have organised the examinations of the workers. According to the opinion of the experts from the MAL, "it should be supported by the industry and the workmen themselves instead of gambling, smoking, alcohol, and other similarly questionable activities." [1] The 4th division planned to cooperate in this task with the Central Association of Industrialists.

The examples of popular scientific booklets created under the auspices of the MAL are: The Calculation of Changeable Gears when Cutting Screws on the Lathe, Tempering, The Introduction to Geodetic Measurements (Mining), The Operation and Linking of Transformers, Displaying Alternating Currents, The Operation of Transformer Station, The Operation of Steam Boilers, Sawing Machines, Wood Bending, Casting, Water Vapour for Machines, Calculating Forgings for Mechanical Blacksmiths, etc...

In that time, the MAL supported also the development of socio-engineering, which among others investigated the influence of socialisation of education on state productivity. The Department of Psychotechnics at the MAL [3] was then focused on the choice of vocation by psychotechnical methods to correspond to the abilities of the individual so that his or her talent would be used in a proper way and the vocation would also enrich his or her life.

2. Conclusion

The publication activities of the 4th division of the MAL were quite broad. Unfortunately, the great thoughts of the MAL were not widely applied with the exception of several popular scientific booklets and courses for workmen. However, the above described ideas concerning the education of wide public supported by the Masaryk Academy of Labour were timeless. Later on, these ideas were substituted by the evening classes for workers at high schools, where workers could develop their knowledge after their working time. Today, there are combined studies at universities, where the students with high school education can study in spite of being employed and thus develop their potential for success in a way that the engineers of the Masaryk Academy of Labour would wish to.

Acknowledgements

Research described in the paper was supervised by Prof. M. Efmertová, Supervisor, FEE CTU in Prague and supported by the Student grant competition of the CTU under grant No. SGS17/131/OHK5/2T/13 - Technical Professions: Technical Engineers and the Czech Society (Between Scientific Milieu, National Prestige and Entrepreneurial Activities) in the Years 1881-1945.

References

Archival materials:

[1] Masaryk Institute and Archives of the AV ČR, fund of the Masaryk Academy of Labour, sign. VII), stock number 881, cardboard 288.

[2] Masaryk Institute and Archives of the AV ČR, fund of the Masaryk Academy of Labour, sign. IX), stock number 953, cardboard 313.

[3] Masaryk Institute and Archives of the AV ČR, fund of the Masaryk Academy of Labour, sign. XIII), stock number 1157, cardboard 362.

Literature:

[4] KUČERA, Rudolf. Nutnost a prostředky demokratisace vzdělání: Organisace mimoškolního studia hlavně technicko-průmyslového. *Časopis československých inženýrův a architektů: Technický obzor: zprávy sdružených spolků československých inženýrův a architektů v Brně, Plzni a Praze*. Praha: Spolek československých inženýrův a architektů, 1921, (29), 167-171; 175-179; 181-182. ISSN 1804-8692.

[5] HOŘEJŠ, Miloš. Studijní dráha ing. Otakara Podhajského, pozdějšího továrníka v Hostivaři. *Dějiny vědy a techniky*. Praha: Národní technické muzeum, 2005, (13), 15-25. ISSN 1804-6622.

[6] HOŘEJŠ, Miloš. Cesty za vzděláním a praxí konce 19. a začátku století 20.: na příkladu pětice podnikatelů ve strojírenském průmyslu. *Dějiny vědy a techniky*. Praha: Národní technické muzeum, 2006, (14), 103-121. ISSN 1804-6622.

Electronic sources:

[7] Prof. Ing. Jan Zvoníček. *Internetová encyklopedie dějin Brna* [online]. [qouted 2019-03-13]. Available at: https://encyklopedie.brna.cz/home-mmb/?acc=profil_osobnosti&load=2309

[8] Prof. Ing. Jan Zvoníček. *Týniště nad Orlicí: Oficiální internetová prezentace města* [online]. [quoted 2019-03-13]. Available at: https://www.tyniste.cz/cs/mesto/historie-mesta/vyznamni-rodaci-a-osobnosti/prof-ing-jan-zvonicek.html

About Authors...

The author **Vit HOLEČEK** was born in 1987 in Hradec Kralove, where he studied at the local university Historical Sciences. His professional focus is Archival Science. He specializes in technical and economic history, particularly in the Czech history from the 2nd half of the 19th century to the 1st half of the 20th century, the history of science and technology and historical development of scientific and technical institutions. Currently he studies PhD studium History of Technology at the Czech Technical University in Prague.

History of Teletext Services in Context of Hearing Impaired People

ZDENEK BUMBALEK

[1] Faculty of Electrical Engineering, Czech Technical University in Prague,
Technická 2, 166 27 Praha, Czech Republic

bumbalek@transkript.cz

Abstract.

The history of teletext is closely connected to the world of deaf people and their need for subtitling TV shows. Teletext was invented in the UK in the early 1970s and its task is to transmit a variety of textual information along with a television signal. The inspiration for the development of teletext was the closed-captioned service for the deaf in the USA, which used a hidden line of television signal to transmit text. Although teletext is one of the most important technical discoveries for the Deaf in the 1970 – 1990 period, it was the Deaf community and their lobby, which by enforcing exclusive use of specific line of TV signal for transmitting closed captions, delayed the introduction of teletext in the US for a decade after European countries.

Keywords

Telematics, teletext, closed captioning, hearing impaired people, history.

1. Subtitles in the Context of Hearing Impaired People

When the first films were created in the first decade of the 20th Century, they did not need any subtitling or closed captioning as they were largely silent films interspersed with information with the help of inter-title cards. This allowed the individuals with hearing impairments to enjoy cinema until the 1920's and 30's. When the first talkies came into existence in the 30's, it ushered an era where deaf people were largely ignored by the entertainment industry until the 1970's and 80's [1].

The first instance of captioning for the deaf occurred in 1958 in the United States of America when the Captioned Films for the Deaf program was signed by President Eisenhower on September 2nd, 1958 [2]. This was the first time the community was specifically tend to after sound was introduced in films in 1927. In 1947, Emerson Romero who was a deaf actor attempted to create the concept of subtitles by acquiring film reels and splicing titles between takes. Although he was not widely successful in his attempts, his efforts led to the creation of the Captioned Films for the Deaf Inc. by Dr. Edmund B. Boatner. He was the superintendent of the American School for the Deaf and his non-profit organization played a major role in pioneering technology for the individuals with hearing impairments [2].

The need for additional assistance for deaf people was recognized by Boatner's efforts, and petitions were sent to Congress in order to shed light on the issue. Previously, the nation had created a Talking Books Program for the Blind in 1933, and this acted as a precedent for the deaf community. This led to the establishment of the Captioned Films for the Deaf Program in 1958. Nearly 80,000 dollars was appropriated for this program in 1959 by the US government and this led rapid strides being taken in order to help the deaf community.

"Although the initial purpose of the Captioned Films program was to provide subtitled Hollywood films for deaf people, educators were quick to recognize the potential of captioned films and other visual media as tremendous untapped educational resources. Consequently, the Congress amended the original law several times. The more significant amendments, for the purpose of this article, were Public Law 87-715 (1962), which authorized research, training, production, acquisition, and the distribution of educational media; and Public Law 89-258 (1965), which authorized the distribution of media equipment." [2]. This led to the corporation being used as a medium in order to design unique educational materials for those with hearing impairments. The technology was furthered in the country over time and it led to the development of closed captioning. In 1976, the FCC issued a decree regulating the technical implementation of closed captions and the exclusive use of the measuring line (line 21 of the TV signal in the US) for subtitling.

2. Origin of Teletext Services

The teletext service became popular around the world during the latter half of the 20th Century, but it had its roots in the United Kingdom. The four-decade old service found its first home in Ceefax – the first teletext service created in the world. When it was launched in Sept., 1974,

the number of users that were aware of the service were highly limited and concentrated in the Great Britain region.

"Today, it's quite easy to talk about the blocky graphics and primary colours of teletext in the past tense, as Ceefax was switched off during the switchover to digital television in 2012. However, it's still running in some countries and its influence can still be felt throughout the internet enabled technology world" [3]. The actual invention of the service, however, was an accident at the time. Engineers at the BBC were working on creating a method with which program subtitles could be offered for those who were hard of hearing. This was done using a portion of the unused section of the 625-line TV signal and the process was dubbed as the field blanking interval.

The creation of this service led to the opening of possibilities related to disseminating information using the television set and field blanking intervals. The same principle was applied in order to display new types of information for the public including news, weather reports, financial data, and sports news. This led to the eventual creation of Ceefax – the first teletext service in the world.

BBC Journalist Colin McIntyre was appointed as the first editor of UK's primary teletext service in 1974. The service was launched in the air a few months after this point. Initially, journalists working with Ceefax were employed with a specific purpose in mind. "Ceefax journalists would monitor incoming wire copy from Reuters, the Press Association and the BBC's own internal news distribution service, GNS, and when a story was to be updated they would type at one of two production terminals and create a Ceefax page" [4].

The service, however, was not compatible with a vast majority of the television sets available to the public during the era. This is why it was accessible only to a very minor section of the population when it was first made available. "When Ceefax launched, it was rumoured that there were only four teletext-capable TV sets in the whole UK - one in the BBC director general's office, one for the director of engineering, one at Kingswood Warren and one in Colin McIntyre's home" [4]. Although the invention of the teletext service was revolutionary and highly relevant for the period in which it was invented, it took the service well over a decade to become a prominent feature in the United States of America. "In the early 1980's three North American television networks provided broadcast teletext services; CBS, NBC, and CBC in Canada. In addition, several other smaller broadcasters, mainly PBS, were involved in providing a teletext service. All that was needed to receive the service was a television set, a decoder, and a remote keypad" [5]. Thus the teletext service managed to break into the U.S and U.K markets with the help of major broadcasting networks.

3. Technical solution

The first major teletext service in the world, CEEFAX, was announced by the BBC in 1972, and by the middle of 1973, test broadcasts were already underway. On 23 September 1974, a 30 page public teletext service was broadcast by the company and this is the period ORACLE also begins full transmission from its centers. Transmission of subtitles was the next item of priority for this technology, and this was achieved in two years when the first ever teletext subtitles were broadcast by CEEFAX. Within the next year, the program managed to acquire numerous upgrades, and this included broadcasts in colored backgrounds, functions to change text size, etc.

By 1978, public interest in the service had begun to peak and numerous citizens around the nation of U.K began to purchase external adaptors for the television set in order to utilize the teletext services. By 1980, televisions sets themselves began undergoing serious modification, and citizens were now able to purchase devices that were already teletext capable. By 1983, there were 1.5 million people with teletext capable television sets in their home.

The middle of the 1980's saw the popularity of teletext services hit its peak, and numerous organizations were looking to join the innovation occurring in the field. By 1987, the concept of teletext add-ons had become a part of the zeitgeist, with consumers being able to utilize various softwares in order to improve their teletext features. The early 90's witnessed some vital changes in the power structures responsible for driving teletext innovation, as Oracle shuts down in 1992. "During the mid-1990s "traditional" teletext settled into comfortable maturity, with the introduction (for CEEFAX) and development (for both services) of regional Teletext services being the most significant change during this time. Telesoftware may have gone, but PCs and add-on teletext cards brought new ways to manipulate data, and the introduction of data broadcasting to teletext introduced new applications" [6].

The change in technology responsible for providing entertainment to the public also drove major changes in teletext technology. The rise of VCR's in the 1990's meant that teletext had to adapt to being utilitarian in the VCR medium. "VCRs began to develop "intelligence", able to read Teletext programme listings and then to automatically set themselves up and start and stop under the control of PDC (which started on Channel 4 in 1991, spread to BBC2 in late 1995 and thence to most other terrestrial channels during the following few years)" [6].

By the end of the 1990's, teletext services were virtually available at any country across the planet. The innovation of digital teletext occurred in 1999 and new information was added to newly available digital television services. Soon, cable companies began to introduce specialized changes to the service with Sky TV creating a new breed of service called "active text". The service continues to be utilized in various countries today and it has changed individually in each country based on the necessity of the people.

4. Comparison of development in UK and France

The United Kingdom pioneered the concept of teletext, but numerous countries managed to duplicate the formula within a decade of its conception. The French created their own version known as Teletel and Minitel in the 1980's and these were popularly used by the audience of the country over the next two decades. "The Télétel/Minitel emerged from French government policies and investments aimed at making France a leader in developing innovative teletext (printed data texts that could be read and interacted with on television and computer screens) that would make databases available to viewers, revolutionize television viewing, enrich the telephone experience and lead to the emergence of an information society"[7]. The service was broken down into two major categories in the nation: Broadcast teletext and Videotex. "Broadcast teletext, composed of data pages transmitted through the airwaves like regular TV broadcasts and accessed through special internal or set-top decoders" [7]. The other major service, Videotex, was an interactive model that managed to connect consumers to various types of large computerized informational databases with the help of the operating telecommunications network.

The French teletext service was highly unique and it managed to add numerous vital changes to the concept of teletext over the course of its evolution. "Broadcast teletext was a revolutionary development because it was a free digital system based on an amalgamation of computers and broadcast technologies" [7]. This system was eventually adopted for a host of purposes including captioning for those with hearing impairments, subtitling of different languages, and presentation of important snippets of information related to the news, sports, and weather.

Developing a successful teletext service in France was largely possible due to the creation of an important service known as Antiope. "One of these innovative technologies was Antiope, a neologism that describes the viewdata phenomenon: Acquisition numérique et télévisualisation d'images organisées en pages d'écriture (Digital Acquisition and Remote Visualization of Images Organized in the form of Written Pages). Antiope was launched in 1972 by the Centre Commun d'Etudes de Télévision et de Télécommunication (Joint Centre for the Study of Television and Telecommunications, CCETT)" [7]. Antiope was utilized to serve numerous important functions including enabling transmission of specific pieces of information to various receivers with the one-way transmission signal and creating an interactive, two-way system where users can exchange valuable data at vital times. The Antiope system was a pioneer device in the process of enabling France to create a functional teletext system.

5. Teletext in Czechoslovakia

A pioneer of teletext in the Czech Republic, or Czechoslovakia at that time, was Czechoslovak Television (ČST), which broadcasted firstly teletext on May 1, 1988. However, it began with the first experiments in 1983. In the first year of teletext broadcasting, it offered more than 300 teletext pages. From the beginning, the information was divided into 4 basic blocks: 1) foreign news, home, sport; 2) block of general or socially needed type information, weather, traffic situation, border crossings, exchange rates; 3) economics, science, technology, ecology and 4) a block of cultural information [8].

The teletext broadcast time coincided with the broadcast time of the 2nd program transmitters, which meant about 14 hours a day at that time. The reason for broadcasting only on the 2nd program was more modern equipment of transmitters. From a technical point of view, it was a 1.5 generations teletext. This generation of teletext enabled the transfer of character sets containing special characters of Central European alphabets, including Czech and Slovak diacritics. However, this choice of technical solution resulted in incompatibility of foreign televisions with the correct display of Czechoslovak teletext due to the lack of support for local diacritics [8].

The availability of suitable televisions, resp. decoder was a major obstacle to the faster onset of teletext in Czechoslovakia. A year after the launch of the teletext broadcast, some 8,500 teletext-supporting TVs were available in Czechoslovakia, of which 1,500 were purchased from abroad in cooperation with the Omnia national enterprise specializing in imports from non-socialist states. Major foreign companies important for Czechoslovak teletext were Phillips Mullard, VG Electronics, Grundig, Thomson and Salora. Phillips Mullard held patents for the production of highly specialized integrated circuits needed for coding and decoding teletext, and the purchase of their components was crucial for the production of Tesla Orava home televisions. Grundig, Thomson and Salora were among the major TV importers in Czechoslovakia. However, their acquisition was possible only in specialized shops TUZEX, where it was not possible to pay Czechoslovak crowns but special tuzex vouchers, so-called bones. Bony could only be officially obtained with mandatory exchange for currency. The acquisition of a television in the late 1980s was a costly matter. The average price of a higher-class model supporting teletext was around CZK 20,000, representing about 6.5 times the average wage in the national economy.

There was a more rapid development of teletext in the 1990s, also as a result of the Velvet Revolution and the related release of foreign trade and imports of foreign televisions. In relation to the hearing impaired community, teletext of greater significance has been achieved since 1992 when the first broadcast of closed captioning for the Deaf occurred on 1 July at 4 pm this year. Thanks to the ASNEP Association of the Deaf, Hard of Hearing and their Friends, in 1997 an Act no 135/1997 on the mandatory

percentage of subtitled programs broadcast by public television was adopted. Several revisions followed, and today, by law, Czech Television is obliged to procure at least 70% of its programs with open or closed captions. For private televisions, this percentage is then lowered to a minimum of 15%.

Closed captions are available in the Czech Republic, similarly as in the UK, via the teletext page 888. A special feature of this site was that they were transmitted in every line of analog TV signal, unlike others. The reason is logical, unlike the other pages, it was not possible to wait for the subtitles and it was necessary to decode the closed caption for each image frame so that the subtitle was transmitted in sync with the TV show. Although the television signal is fully digital today, the principle of the closed caption availability via teletext page 888 remains unchanged. Furthermore, there is growing pressure to increase the number of subtitles in both public and private television channels. Newly, the obligation to introduce subtitling for non-linear broadcasting is also discussed, e.g. for Internet TV. This issue is discussed in more detail in the European Parliament and Council Directive 2018/1808, which is to be implemented in Czech legislation by 2021.

6. Concept – knowledge transition and vice versa

When applying the C-K Theory [9] in this study, we see that there are numerous interesting concepts that can be derived from teletext services. One major concept that comes to mind is the idea that important information needs to be easily shared to the public. The concept of breaking news has become a boring façade in today's media frenzy climate, but nearly four or five decades ago, transmitting information to all corners of a country was a near impossible task. More importantly, only a radio could be used to send information, and it relied on the public being open to accessing the information while it was being broadcast.

The concept of emergency news being immediately broadcast to the public has always been a necessity in the world, but the solution for this issue was greatly aided by teletext services. With the help of this service, consumers with their television set on could access important information that was broadcast on a minor space at the bottom of the screen – the concept that eventually came to be known as the ticker.

The conversion of concept to knowledge happened extremely quickly with the establishment of the teletext service. The need for a service where information could be freely, quickly, and promptly dispersed among the public has been a necessity for millennia, but one particular type of technological advancement has allowed for this requirement to be capably fulfilled. With the advent of teletext services, countries around the world were able to create specialized content for their citizens based on the need of the hour. A tornado moving towards one part of the nation could be easily avoided due to the resourcefulness of tele texting capabilities. Thus we see that concept has managed to become knowledge with regard to the concept of emergency broadcasting turning into ticker feeds.

Acknowledgements

Research described in the paper is a direct continuation of the research linked to the supported and already completed grant No. SGS15/178/OHK5/2T/32: Historical contribution of electrical engineering and telecommunications and their professional organizations to improvement of quality of life of hearing impaired people.

References

[1] MAGER, W.: "See Hear: Deaf television past and present". BBC, 2014, https://www.bbc.com/news/blogs-ouch-26629727. Accessed 1 Mar. 2019

[2] NORWOOD, M. J. "Captioning for Deaf People: An Historical Overview". Described and Captioned Media Program, 2019, https://dcmp.org/learn/80-captioning-for-deaf-people-an-historical-overview. Accessed 1 Mar. 2019.

[3] MERRIMAN, C.. "Teletext is 40 Years Old". The Inquirer, 2014, https://www.theinquirer.net/inquirer/news/2372556/teletext-is-40-years-old. Accessed on 27 Feb. 2019

[4] "Ceefax: The Early Days". BBC, 2012, https://www.bbc.com/news/magazine-20032531. Accessed on 28 Feb. 2019.

[5] GRAZIPLENE, L. R. "Teletext: A New Idea for News Dissemination". Teletext: Its Promise and Demise, Lehigh University Press, 2000.

[6] COOK, M.., BROWN, M.: "Teletext Timeline". The Teletext Museum, http://teletext.mb21.co.uk/timeline/. Accessed 28 Feb. 2019.

[7] EKO, L. "Teletext and Videotex in France: From Innovative Social Media to Objects of Cultural Memory". 2016, https://www.researchgate.net/publication/296705257_Teletext_and_Videotex_in_France_From_Innovative_Social_Media_to_Objects_of_Cultural_Memory.

[8] "Teletext dnes a zítra". CST, 1989. https://www.ceskatelevize.cz/ivysilani/676174-teletext-dnes-a-zitra/. Accessed 28 Feb. 2019.

[9] BUMBÁLEK, Z. "Application of C-K Theory within the Historical Development of Means of Communication for Hearing Impaired People" In *Proceedings of the 9th International Student Conference on Electrical Engineering POSTER 2015.* Prague (Czech Republic), 2015

About Authors...

Zdenek BUMBALEK received his master degree in Telecommunication engineering from FEE CTU in Prague in 2008 and 2nd master degree in Management and economy in 2011. He is currently a Ph.D. candidate of at CTU in Prague. His current research interests include history of communications means of hearing impaired people and using modern communication systems in assistive technologies.

The development of DNA microarrays

Vladimír Kunc[1]

[1]Dept. of Computer Science, Czech Technical University, Karlovo náměstí 13, 121 35 Praha 2, Czech Republic

kuncvlad@fel.cvut.cz

Abstract. *DNA microarrays played evolutionary role in the genomics research in several last decades. The DNA microarrays are a method for measuring relative concentrations of nucleic acids. It works by placing a high number of probes consisting of nucleic acids with specific sequences placed on a surface and using hybridization with fluorescently labeled complementary sequences to the probes that are then detected. The purpose of this paper is tho briefly describes the development of the DNA microarrays that occurred in last three decades.*

Keywords

DNA microarray, gene expression, DNA chips, genomics, genotyping

1. Introduction

Two last decades in genomics were characterised by the massive use of oligonucleotide and DNA microarrays which allows to obtain genome-wide mRNA expression data [3]. While the very predecessor of microarrays was used already in 1975 [1, 3], the microarrays are still actively use for research even today (e.g., [2, 4, 5, 7]).

2. Invention of microarrays

The first description of the DNA in 1953 by Watson and Crick [10] lead to the emergency of the genomics. The first steps towards the creation of the DNA microarrays were presented in the late 60s when the way for locating the position of specific sequences (*in situ* hybridization) was discovered [8]. The method became known as FISH (fluorescence *in situ hybridization*) after introduction of fluorescent probes [8]. The FISH uses fluorescent probes that bind only sequences with a high degree of complementarity which can be then observed using a fluorescence microscopy.

2.1. First arrays

The development of colony hybridization method lead to the creation of first microarrays; the colony hybridization method randomly cloned the probe DNA into *E. coli* plasmids, and then grown the colonies thus replicating the probe DNA. This was used for first larger scale experiments where it was used to screen thousands of colonies to identify clones with the DNA complement to the probe DNA [1]. The approach was later extended in 1979 by creating an array of 1728 colonies in a 26×38 cm region [1]. The extension was done by creation of a mechanical pin devices that allowed to operated simultaneously on 144 well microplates. The first and simplest arrays were the called *dot blots* with simplified processing and better reproducibility — they allowed for parallel hybridization and also for parallel image processing [8]. The density of the first array was further increased by replacing the manual work by robotic systems which also removed the humans errors that inevitably occurred [8].

2.2. Increasing the density

The *dot blot* procedure used a porous support as it provided a larger surface for binding and also the nucleic acid could be applied in relatively large volumes because it soaked into the porous material and thus preventing excessive lateral spreading [8]. The porous support came, however, with several disadvantages — the boundaries and shapes of spots were poorly defined and it was difficult to control the amount deposited oligonucleotide. The porous support was an obstacle in increasing the density of spots which was necessary to increase the number of spots. Furthermore, the permeable membranes tended to swell in solvent and to shrink and distort when dried and also the nonrigidness made spotting and reading their position more difficult [6, 8].

The solution was to replace the porous support with an impermeable materials such as glass or silicon which allowed the use of very small sample volumes and high density of spots [6]. Furthermore, since the nucleic acids form a monolayer that saturates the surfaces, the impermeable support allowed for the consistency of attached amounts between regions of the array [8]. Also, such supports increased the reaction speeds of the solution phase as the molecules did not have to diffuse into and out of the spores [8]. The impermeable supports allowed the technology to reach high accuracy, reliability and reproducibility needed for larger scale experiments.

2.3. Modern microarrays

After the introduction of impermeable supports, three major direction of microarrays emerged — spotted arrays, in-situ synthesised arrays, and self assembled arrays [1].

Spotted arrays

The first method allowing high density arrays on glass substrates was published in 1996; the method used poly-lysine coated glass microscope slides that provided good binding of DNA and also a robotic spotter that was spotting multiple glass slide arrays from DNA stored in microtiter dishes [1].

In-situ synthesised arrays

Another direction represented the *in-situ* synthesis of nucleic acid on the surfaces which brought multiple advantages over deposition of presynthesized probes [8]. One of possible approaches was the ink-jet fabrication that used ink-jet printers to fire solution of nucleotide reagents at a glass surface. Due to the similarity of ink-jet printing on paper, most of the engineering work was already researched which allowed quick introduction of the method as it required only the modification from printing four colours to to delivering precursors for four different bases [8]. This increased the flexibility of microarrays as it could synthesise any set of oligonucleotides and place it at any position in the array [8].

Different approach is the *light-directed* fabrication (manufactured by Affymetrix) that directed the synthesis of oligonucleotides by using patterned photolitographic masks [1, 8]. A single mask was required for each base base addition; thus to create a probes of 20 nucleotides, 80 masks were required [8]. The biggest advantage of such approach is the high density — there are microarrays that have 65,536 probes in 1.28×1.28 cm area [8].

Self assembled arrays

Yet another method was introduced in 1998; it synthesises DNA on small polystryrene beads and then deposits the beads on the end of a fiber optic array[8]. A randomly assembled array is then created by applying a mixture of such beads to the optic fibre. First versions of the technology used optically encoded beads with different fluorophore combinations to determine the position of the nucleotides. However, this rather limited the number of unique beads that could be distinguished [8]. This was solved in 2004 by decoding the beads using hybridization and detection of a number of short and fluorescently labeled oligonucleotides which allowed for large number of types of beads on a single array and for functionality tests of such array prior to its use in a biological assay [8].

3. Future of microarrays

The DNA microarrays characterized a whole phase in the genomics research. Even though they are still actively used in various researches, a new competitor has arisen recently — it is called the RNA-Seq. It has one great advantage compared to microarrays as the microarrays need to known the analysed sequences *a priori* to the experiment [9]. The RNA-Seq represents a whole transcriptome shotgun sequencing and as such produces reads of sequences in the analysed sample without the need to know the sequences before the experiment — the measurement of the gene expression is done by counting the reads during the transcriptome assembly.

However, the microarrays are often preferred in certain applications as they are cheaper. Furthermore, there are microarray-based tests with proven clinical utility and they are easier to use for diagnosis.

Acknowledgements

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. Additional computational resources were supplied by the Ministry of Education, Youth and Sports of the Czech Republic under the Projects CESNET (Project No. LM2015042) and CERIT-Scientific Cloud (Project No. LM2015085) provided within the program Projects of Large Research, Development and Innovations Infrastructures. This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS17/189/OHK3/3T/13.

References

[1] R. Bumgarner. “Overview of DNA microarrays: types, applications, and their future”. In: *Curr Protoc Mol Biol* Chapter 22 (Jan. 2013), Unit 22.1.

[2] D. Edsgärd et al. “Identification of spatial expression trends in single-cell gene expression data”. In: *Nature Methods* 15.5 (Mar. 2018), pp. 339–342.

[3] A. A. Ewis, Z. Zhelev, R. Bakalova, S. Fukuoka, Y. Shinohara, M. Ishikawa, and Y. Baba. “A history of microarrays in biomedicine”. In: *Expert Review of Molecular Diagnostics* 5.3 (May 2005), pp. 315–328.

[4] M. Hao, B. Barlogie, G. Tricot, L. Liu, L. Qiu, J. D. Shaughnessy, and F. Zhan. “Gene Expression Profiling Reveals Aberrant T-cell Marker Expression on Tumor Cells of Waldenström’s Macroglobulinemia”. In: *Clinical Cancer Research* 25.1 (Oct. 2018), pp. 201–209.

[5] A. Y. Higashi, B. J. Aronow, and G. R. Dressler. “Expression Profiling of Fibroblasts in Chronic and Acute Disease Models Reveals Novel Pathways in Kidney Fibrosis”. In: *Journal of the American Society of Nephrology* 30.1 (Dec. 2018), pp. 80–94.

[6] T. Lenoir and E. Giannella. “The emergence and diffusion of DNA microarray technology”. In: *Journal of Biomedical Discovery and Collaboration* 1.1 (2006), p. 11.

[7] J. B. Nielsen et al. “Biobank-driven genomic discovery yields new insight into atrial fibrillation biology”. In: *Nature Genetics* 50.9 (July 2018), pp. 1234–1239.

[8] E. M. Southern. “DNA Microarrays: History and Overview”. In: *DNA Arrays*. Humana Press, 2001, pp. 1–15.

[9] Z. Wang, M. Gerstein, and M. Snyder. “RNA-Seq: a revolutionary tool for transcriptomics”. In: *Nature Reviews Genetics* 10.1 (Jan. 2009), pp. 57–63.

[10] J. D. Watson and F. H. C. Crick. “Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid”. In: *Nature* 171.4356 (Apr. 1953), pp. 737–738.

Personal sources relating to the electrotechnical education at VŠB

Mariana STONIŠOVÁ[1]

[1]Czech Technical University in Prague, Jugoslávských partyzánů 1580/3, 160 00 Prague 6, Czech Republic

stonimar@cvut.cz, m.stonisova@seznam.cz

Abstract. *In the paper, which is based on the upcoming dissertation, I focus a great attention on university technical intelligence as a bearer of scientific branch. A basis of the thesis is an archive research and a critical source analysis. I deal with particular chosen personal sources, i.e. videos and diary records belonging to the professors from the electrotechnical branch at VŠB. The first part of the text is devoted to the sources classification and only then I focus on particular personal sources, their origin, characteristics and what qualitative value do they have for the particular topic. As examples I chose video recordings with Tomáš Čermák and personal diaries written by Vladimír Chůra.*

Keywords

Čermák, Chůra, Personal diaries, Sources, Video record,

1. Introduction (reasons for the choice of the sources)

In the dissertation, I focus a great attention on university technical intelligence as a bearer of scientific branch.[6] The main concern is devoted to professors who led the electrotechnical institute (department) and schoolmasters who had merit in contribution to the scientific educational system within their branch. I create their biograms by means of a microhistory and biography method. Thus, I can consider the schoolmaster as an individual in various contexts (an independent personality as well as a part of an entity – family, society, employee etc.). The scientist`s personality plays an unsubstitutable role for the particular branch because it is the personality that creates him in his entire complexity.[2] By means of a synthesis and based on created biograms, scientific profiles and on their scientific-research activities, I can point to development of electrotechnical fields at VŠB. According to Vaněk (Vaněk, p. 85), a method of oral history is successfully used from an individual`s biography to development of an institution. It is a part of this article as well.[11]

A basis of the thesis is an archive research and a critical source analysis. In the article, I deal with particular chosen personal sources, i.e. videos and diary records belonging to the professors from the electrotechnical branch at VŠB. It is because the personality belonging either to the schoolmaster or to the scientist forms the branch and its surroundings as well as he/she himself/herself is influenced by the environment where he/she operates and by social context of the particular period. The first part of the text is devoted to the sources classification and only then I focus on particular personal sources, their origin, characteristics and what qualitative value do they have for the particular topic.

As examples I chose video recordings with prof. Tomáš Čermák and personal diaries written by prof. Vladimír Chůra. I focused mainly on Tomáš Čermák as he spent several decades (from the 1960s to the first decades of the 21st century) at VŠB either as a student or as a scientific worker. Vladimír Chůra`s personal diaries belong to sources which I have not published so far and they were not utilised too much on a scientific level.[7, 9, 10]

2. Classification of sources

When classifying sources, we apply various viewpoints: outer intention, price of sources for historical knowledge, relationship to individual parts of human action. Within the thesis, I stick to the traditional division of historical sources according to the way of their preservation (in accordance with outer intention). This division includes material, orally preserved, written (authoritative and private character, narrative sources) and visual sources. In each presented category, we could find personal sources.[4, 5, 11] However, as in the article I intend to focus on videos with VŠB schoolmasters and their diaries, the following part will introduce these orally preserved and written sources.

3. Orally preserved sources on a video record

For a historian, contemporaries`narration concerning various historical events means an invaluable orally preserved source. These events are during narration written down or recorded.[4] According to Vaněk, this personal source belong to main forms of utilization of the method of oral history for a historian. We speak about a life story (narration) where there is an obvious attempt to help clarify a certain historical process or period (limited by the narrator`s actual life) from the narrator`s point of view or directly his destiny, experience and standpoints. In the same way as the interview *"is conveyed with an intention to enlarge and deepen historical knowledge of the particular period, event or topic."* (Vaněk, p. 83).[11] We focus on narrator`s direct experience during the interview/narration. Video recordings allow the narrator to monitor non-verbal communication. We have to take into consideration.[11]

The source contains many insufficiencies and it has to be scarified. The narrator`s personality presents the greatest problem as it is influenced by memory, observational abilities, education and another aspects. The contemporary was a part of the school events he refers to, however, we have to keep in view what position does he present within the described event and if he distorts the utterance or not (defence of deeds etc.). Furthermore, we have to bear in mind that the memories were recorded in the course of time (narration about past is carried out in present) and that they could be distorted by later interpretation (social and political situation, medial version of reality). As the video record sometimes replaces missing messages about events and about background of school events, a comparison with other sources is important.[4, 5, 7, 11]

3.1 History of video records origin

In 1999, the VŠB-TUO archive employees in cooperation with VŠB-TUO Audio-visual Centre started to record videos with biographical narration of schoolmasters, scientific workers and graduates belonging to this school. During the first phase, graduates who experienced studies at VŠB in Příbram were addressed. They could record school atmosphere in the First republic period, local schoolmasters and moving of the school from Příbram to Ostrava. Since 2010, bearers of scientific branches have been recorded.[10] The addressed individuals included also those who experienced blossom of electrotechnical branches from the sixties in the twentieth century as well as difficulties that they had to encounter during establishment of the electrotechnical faculty. These include prof. Ing. Tomáš Čermák, CSc., prof. Ing. Ivo Vondrák, CSc. a prof. Ing. Vilém Srovnal, CSc. For the time being, these three records are available in a "crude" form in the VŠB-TUO archive. However, there is a plan that they will undergo further editorial revision. In the article, I focused mainly on Tomáš Čermák as he spent several decades at VŠB either as a student or as a scientific worker. The narrator remembers his personal life (parents, education) and mainly the scientific branch which he specialized in and which he accompanied during its birth. Furthermore, he thinks about cooperation with industrial practise. The memories depict social and political atmosphere at school, eventually in our state.

3.2 The video record with prof. Ing. Tomáš Čermák, CSc. and informative capability for the particular research/chosen research theme[1, 3, 10]

The video record with Tomáš Čermák was made in 2011 and it is possible to divide it into two parts with approximately same length (2x45 minutes). In the first part, he speaks about his personal life, studies, scientific practice in industry, scientific-pedagogical activities at VŠB and to his colleagues and students. In the second part, he is asked by the archive worker leading the interview to comment on anything from 1989 as he was working at school at that time. The contribution focuses on the first part as it is devoted to the electrotechnics.

The narration starts with information about family relations and the basic as well as the high school. It is followed by entering the university and his subsequent shift to electric drives. He does not forget to emphasize meaning of his choice with regard to contribution for developing industry that he enters after graduation from the university. In 1968, he started to work as a technical assistant on the Electrotechnical Department of the Faculty of Mechanical engineering where he made use of his experience from practice. He remembers his superiors (e.g. prof. Vladimír Chůra) and their contribution for science and technique. Even though the department schoolmasters provided education of electrotechnical subjects for all branches, prof. Čermák and his colleagues were specialized from the branch point of view. T. Čermák specialized in electric machines and drives for which he compiled several lecture notes.

The 1960s brought enlargement of the existing subjects with new ones. The lectures were innovated and laboratories were built, as they had been insufficient by then. T. Čermák labels them as a technical museum of everything that was not needed by the industry anymore. However, such laboratories brought no contribution for students` development of mind. The new laboratories were located in Slezská Ostrava.

In 1960s, the grounds for construction of laboratories in Poruba together with a project for construction of new school premises in Ostrava Poruba were being prepared. Prof. Čermák was in charge as a deputy of the Electrotechnical Department and a secretary. Already then, there was an attempt to create a technical university from the mining school and to develop an electrotechnical faculty. The attempts had not been successful so far. There was an argument against establishment of the faculty noting that *"you cannot have an electrofaculty because you*

do not have professors and you cannot have professors because you do not have a faculty".[10] At the turn of 1969/1970 the first independent study branch called high-current electrotechnics was opened. It was established due to a lack of high-current engineers in industrial companies in the Northern Moravia District. Full-valued studies had to be secured with minimum number of docents and professors so at the beginning, external workers were commuting from Brno and Prague to Ostrava. Young people who had finished their post-graduation replaced the external workers in the course of time. T. Čermák was in charge of providing the external workers.

The Electrotechnical Department provided education of all electrotechnical subjects within the new branch until 1977. The department was subsequently divided into the theoretical and general electrotechnics department, the department of electrical machines and drives and the department of electricity industry. At the same time, the name of the Faculty of Mechanical Engineering was changed into the Faculty of Mechanical Engineering and Electrotechnics. During this period, there was an obvious strong contact between the school and industrial practice. Apart from the fact that within teaching the professors could utilize their previous practical experience gained outside the VŠB, they also applied science into practice. The industry supplied the faculty with topics for grade projects and diploma theses. They also cooperated with Vítkovice, Nová Huť etc. Contacts with foreign schools (Dresden, Chemnitz) were kept. T. Čermák had an opportunity to develop these contacts on a personal level.

In 1980s, a new branch of microtechnology was established after reorganization of departments (the department of electrotechnics and electric drives and the department of electric machines were set). When prof. Čermák became a sub-dean of the Faculty of Mechanical Engineering and Electrotechnics in 1985, he tried to establish a new faculty. As a member of the commission of experts, which prepared nationwide study programmes, he had an insight into questions of an optimal structure of study programmes. He had merit in gaining a right to defend candidate works as well as in equipping the school with computers.

After the department reorganization, the Department of electronics and electric drives led by Tomáš Čermák (who was a docent at that time) was established.

At the beginning of the nineties of the twentieth century, the Department of Microelectronics, Measurement and Regulation technique and the Department of Informatics were established. Further experts entered the university so at the date of the new faculty establishment there was a sufficient number of schoolmasters who could provide education at the new faculty. In 1990, VŠB asked the Ministry of Education for establishment of a new Faculty of Electrotechnics. Its establishment was confirmed on 1 January 1991. On 1 September 1993, its name was changed to the Faculty of Electrotechnics and Informatics.

Fig. 1. Prof. Ing. Tomáš ČERMÁK, CSc., dr.h.c. (Reference: BIOLKOVÁ, J., KAŠING, P. *Direktoři a Rektoři Vysoké Školy Báňské – Technické univerzity Ostrava*. Ostrava 2014.)

4. Written personal sources – personal diaries

The group of written personal sources includes sources that are connected by their rise with public and private life of individuals (they are private by their origin and they provide evidence about public events) or eventually their families.[4] They can be found mainly in inheritance or personal funds. It includes private correspondence, diaries and journals that did not aim to provide evidence about life publically, as well as memories with autobiographies that were supposed to inform further generations. All of them are of a subjective character and express a personal standpoint, which the author could not or did not want to convey publically.[5]

According to Vaněk the method of oral history is connected with various ways of history record that consider an individual`s conveyance an authentic historical source. Such source can mean memories, diaries, personal letters, school compositions etc. that capture individual feelings, experience and ideas but lack a necessary criterion of oral history: *"by a method of a one to one interview between a questioner and a narrator. A written conveyance which is "not revised" and "not censored" follows already during the process of writing itself other regularities than an utterance recorded on a sound (eventually audio-visual) carrier. Even though absence of the second person (a questioner) can prevent the writer from hesitation, it does not expose him to necessary confrontation with recording technique and it does not evoke an impression of "a public speech", the written formulation of ideas, experience and feelings and opinions is a different way of conveyance from the qualitative point of view."*(Vaněk, p. 75) [11]

In the following part I focus on Vladimír Chůra`s personal diaries because they belong to sources which I have not published so far and they were not utilised too much on a scientific level.

4.1 Characteristics of diaries written by prof. Ing. Vladimíra Chůra, CSc. [9]

The professor Vladimír Chůra was writing personal diaries during the course of his life. The diaries contain precious information about happening at VŠB and are owned by V. Chůra`s widow, Světlana Chůrová. Due to the fact that the diaries contain records depicting not only the work part of Chůra`s life but also his personal life, Mrs. Světlana Chůrová provided the VŠB-TUO archive only with extracts from the diaries relating to her husband`s professional life and happening at VŠB, in 2017. On one hand, we have at least some of his notes, on the other hand, the notes were taken out of context and we cannot objectively judge if there were any mistakes made during transcription or if the past was not changed etc. Therefore, the source has to be deeply examined. As within his diaries, professor Chůra mentions also persons who are still alive, we have to bear in mind protection of privacy and personal data.

4.2 Personal diaries written by prof. Ing. Vladimír Chůra, CSc. and informative capability for the particular research/chosen research topic[1, 3, 8, 9]

In 1950s, significant experts (including Vladimír Chůra) from practice entered the school. They had professional contacts with operational practice. The Institute of Electrotechnics disaffiliated. It was attributed to the Faculty of Mining Machinery as a part of the Department of Machinery. Construction of electrotechnical laboratories was started. Due to parallel costly building of the Czechoslovak Academy of Sciences, whose institutes provided scientific research in preference, the research activities at universities could not be sufficiently secured from the material, equipment and personal point of view. The independent Department of Electrotechnics was founded in 1959. Vladimír Chůra led the department between 1958 and 1969 and he wrote the nationwide textbook *Electrotechnics for mining and metallurgical engineers* (1972). He states reasons for this writing in the diary. He was interested mainly in metallurgical electrotechnics. Between 1969 and 1972, he lectured the subject Metallurgical electrotechnics at the Technical Army Academy in Cairo.

In his diaries (respectively in their extracts), professor Chůra presents his relation to happening at school as well as his relationship to particular people mainly at the turn of 1950s and 1960s that were influenced by political events. He describes in detail his scientific and pedagogical activities not only at VŠB in Ostrava but also for example in Cairo (after some time here because as he says, he did not have time to write the diary). Another remarkable fact is his return to VŠB during celebration of his 70th birthday in the nineties as well as evaluation of local electrotechnical laboratories or an overall shift of the branch to which he devoted his time.

Fig. 2. Prof. Ing. Vladimír Chůra, CSc. (Reference: University Archive of the VŠB-TUO. The Fund: VŠB-TUO, The Rectorate, The Personal Files, V. Chůra.)

5. Conclusion

In this paper, I dealt with special personal sources – video recordings and personal diaries of VŠB professors, in particular with the life story of prof. Tomáš Čermák on video and personal diaries of prof. Vladimír Chůra. Both sources provide valuable information not only about the private life of professors, but mainly about the professional activities at VŠB and about the development of electrical engineering at VŠB. They complement the existing literature and sources of official character. Both sources describe the evants at VŠB in the 1960s to 1990s. Thanks to these sources, we can better understand the socio-political context of this time. Working with these sources requires careful criticism. In another research I count on using video recordings with professors Vondrák and Srovnal from VŠB and with editorial transcription of recordings.

Acknowledgements

Research presented in this paper was conducted under the supervision of employees of Archives of VŠB-TU of Ostrava and it was supervised by Prof. PhDr. Z. Beneš, CSc., Faculty of Arts Charles University in Prague and supported by grant No. SGS17/131/OHK/2T/13. Thanks also to the Faculty of Electrical Engineering of the CTU in Prague.

References

[1] BIOLKOVÁ, J., KAŠING, P. *Direktoři a Rektoři Vysoké Školy Báňské – Technické univerzity Ostrava*. Ostrava 2014.

[2] DOSTÁL, D. *Osobnost vědeckého pracovníka*. Olomouc: Univerzita Palackého v Olomouci, 2013. ISBN 978-80-244-3319-6.

[3] *Fakulta elektrotechniky a informatiky. Vysoká škola báňská – Tehcnická univerzita Ostrava.* Ostrava 1996.

[4] HAVRÁNEK, J., MYŠKA, M. and PACLÍK, J. *Úvod do studia dějepisu.* Praha : Státní pedagogické nakladatelství, 1967, 209 pp.

[5] KRANDŽALOV, D. ed. *Základy studia dějepisu*. Praha: Státní pedagogické nakladatelství, 1964. 346 s.

[6] MYŠKA, M. *Podíl technické inteligence na humanizaci ostravské průmyslové oblasti v počátcích moderní industrializace.* In: Humanizace technických věd: sborník z mezinárodní vědecké konference, sekce 36. Ostrava 1995, s. 59–62.

[7] STONIŠOVÁ, Mariana. *Problematic Aspects of Archival Research on the Development of Teaching Electrical Engineering Using the Example of Leading Personalities this Field of the VŠB-Technical University of Ostrava in the Czech Land.* In Lecture Notes in Management Science. Volume 91. Singapore : Singapore Management & Sports Science Institute, 2018, s. 118–123. ISSN: 2251-3051. ISBN: 978-981-11-7596-1 (Konference: ESE International Conference on Sports, Health and Management (ESE-SHM 2018) (11.–13.6. 2018; 8. : 2018 : Paříž, Francie))

[8] University Archive of the VŠB-TUO. The Fund: VŠB-TUO, Faculty of Mechanical and Electrical Engineering, The file of habilitation and professorship proceedings, V. Chůra.

[9] University Archive of the VŠB-TUO. The Fund: VŠB-TUO, The Rectorate, The Personal Files, V. Chůra.

[10] University Archive of the VŠB-TUO. The video records with Čermák, 28. 1. 2011; Santarius, 6. 11. 2010; Srovnal, 25. 2. 2011, Vondrák, 23. 7. 2010.

[11] VANĚK, M., MÜCKE, P. and PELIKÁNOVÁ, H. *Naslouchat hlasům paměti: teoretické a praktické aspekty orální historie*. Praha: Ústav pro soudobé dějiny AV ČR, 2007. 224 s. ISBN 978-80-7285-089-1.

About Author

Mariana STONIŠOVÁ is a Ph.D. student of the all-school doctoral study program History of Technology at The Czech Technical University in Prague. She works as an academic staff member at Department of Social Sciences under the VŠB-Technical University of Ostrava. She was a member at Department of Social Sciences under the VŠB-Technical University of Ostrava. She was a member of authors collective who published Ostravské dělnické kolonie I., II., III. (2011, 2012, 2015). She cooperates with the archives of VŠB-Technical University of Ostrava on publications about the history of this university.

History of transmissions in Mladá Boleslav

Lukáš NACHTMANN

[1] ŠKODA AUTO Corporate Historical Archives, Tř. Václava Klementa 294, 293 01 Mladá Boleslav, Czech Republic
[2] Dept. for Science and Research, CTU (OVV ČVUT pro CDSP HT), Jugoslávských partyzánů 1580/3, 160 00 Prague 6-Dejvice, Czech Republic

lukas.nachtmann@skoda-auto.cz

Abstract. *Automaker Škoda is one of the oldest car manufacturers in the world. For its current production it uses the latest trends and technology, for example in the case of transmission it uses hardening of pressurised helium. A bonus that automaker Škoda has is its tradition, along with awareness that the future has roots in the past. Therefore, it is necessary to realise that as the inter-generational handover of technical know-how occurs, it is necessary to capture and process information not only for the next generation of skilled hands, but also for the wider technical community. In the past 2 ½ years my colleagues from the technical development department of the automaker and I have worked on the preparation of a publication on the history of the transmissions developed, produced and used by the famous Czech car maker.*

Keywords

Technical history, transmissions, gearboxes, Czechoslovakia, Laurin & Klement, Škoda, automotive industry

1. Introduction

Car maker Škoda has its headquarters in Mladá Boleslav, where it was founded 124 years ago. Václav Laurin and Václav Klement were with it at the beginning. In 1895 they began repairing and producing bicycles. Much has changed since then, and the cars' coolers do not have the L&K logo, but a winged arrow, and the company with quite a few employees is nowadays a global firm. Its products are sold in more than 100 countries. At a time of changing trends, when the use of electric vehicles is in the rise, cars will be manufactured which need a transmission, because it is essential. Today the car maker itself manufactures modern automatic two-gear transmissions (DSG, dual-shift gearboxes) of a single type, but it has also taken over responsibility for all manual transmission in the entire Volkswagen Group

2. History of transmissions

The fact that any use of energy must be optimised was recognised already by our ancestors, when they transitioned from sliding (sleds) to rolling (wheels). The problem with cogged wheels and experiments with transmissions was focused on renowned artist Leonardo da Vinci. The first new-age steps toward a transmission were then made by James Watt during improvement of his steam machine (1796). It took another one hundred years before additional inventors moved transmission technology forward, and they included automotive industry pioneers automobile Carl Benz and Gottlieb Daimler at the end of the 19th century. At the same time, Laurin and Klement of Mladá Boleslav introduced their bicycles onto the market, and they were powered most often by chains, as is common today, or by shafts. In both cases, it is necessary to have a geared wheel or wheelset, and from there it is only a step forward to transmissions.

3. What exactly is a transmission?

A transmission is a relatively overlooked device, which helps a car with what is expected from it: to go slow or fast, up a hill or along a flat surface, or to back up. However, it also applies or actually decides about dynamics of driving, consumption and noise, as well as about the reliability of the vehicle. With a manual transmission, the driver decides when to shift the gears depending on the required speed and switches gears with the gear lever or newly a cable and moves the geared wheel inside the gearbox.

Manual transmissions are usually double-shaft or triple-shaft. One shaft propels, and the other is propelled. Each of them has geared wheels. The geared wheels on the propelling and propelled shafts form a pair, which fit together. Each pair has a different size, and a different number of teeth. One pair of wheels represents one speed level. The principle of connecting wheels and their movement, shifting of gears, has changed over time. Originally entire pieces were moved in relation to each other, and the teeth of the opposite wheels fit together.

However, as of the 1930s wheels with diagonal teeth began being used more, which provided advantages such as noise reduction. However, wheels with diagonal teeth can no longer be moved directly opposite each other.

The quality of shaft material and in particular the material of geared wheels has increased, and the quality of machining has also been refined. It has also been necessary to resolve lubrication, so that parts do not get worn. The original transmissions at the beginning of the 20th century consisted of a few dozen parts, and today's automatic transmission DQ 200 produced in the branch factory in Vrchlabí has 297.

4. Conclusion

During its existence, Škoda has produced hundreds of types of cars. Particularly early on, each type was an original, and each had a different transmission. Why there has not been greater unification is not known. Later, a few types and generations of models used a single type of transmission. For a non-expert this can be considered somewhat confusing, and some cars that look the same sometimes have different transmissions (Š 1000MB X Š 1100MB), while vehicles generationally and visually different have the same transmission Š 1100MB - Š 100).

A major breakthrough in the car maker's history was also apparent in technology, and the introduction of new car designs always went hand in hand with a new transmission, since knowledge and approaches of prior developers were used over a span of 20-30 years.

Fig. 1. Voiturette A (1905-1907)

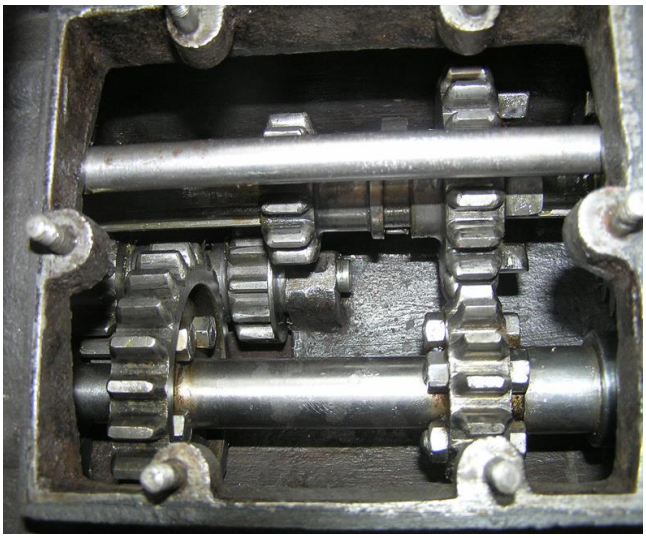


Fig. 2. Voiturette A (1905-1907), transmission - photo

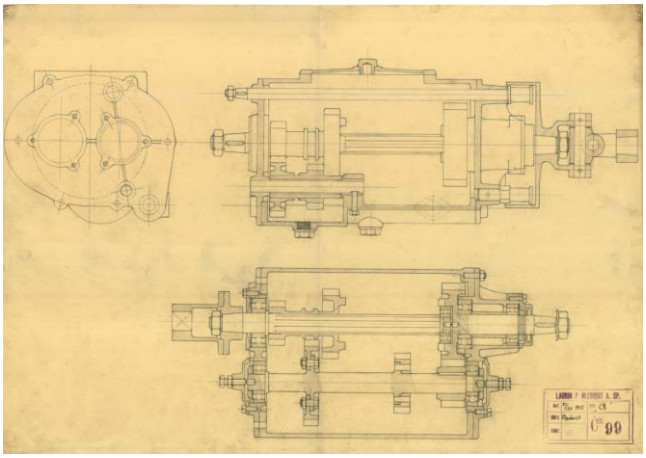


Fig. 3. Voiturette A (1905-1907), transmission - technical drawing

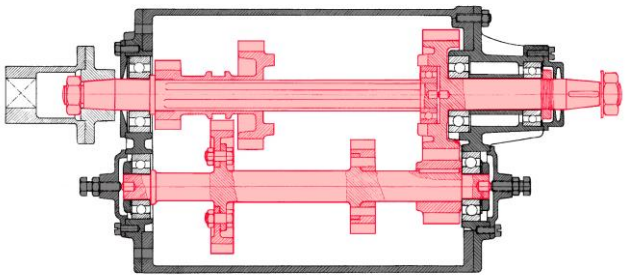


Fig. 4. Voiturette A (1905-1907), transmission - technical drawing, digital midified

Fig. 5. Š 1000 MB (1964-1969)

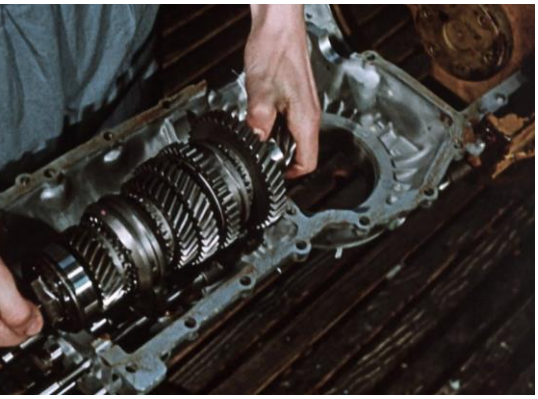


Fig. 6. Š 1000 MB (1964-1969), transmission – photo

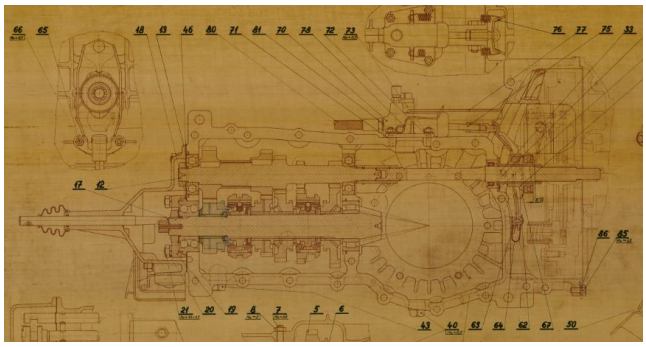


Fig. 7. Š 1000 MB (1964-1969), transmission - technical drawing

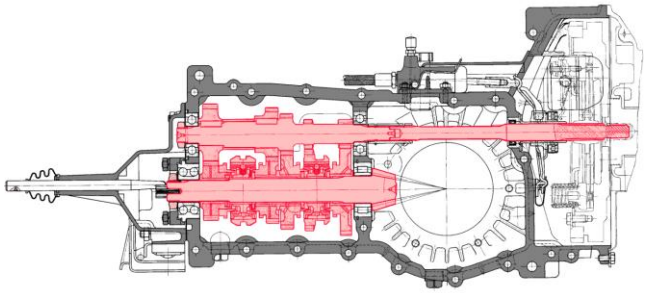


Fig. 8. Š 1000 MB (1964-1969), transmission - technical drawing, digital modified

Fig. 9. Š Favorit (1987-1995)

Fig. 10. Š Favorit (1987-1995), RTG

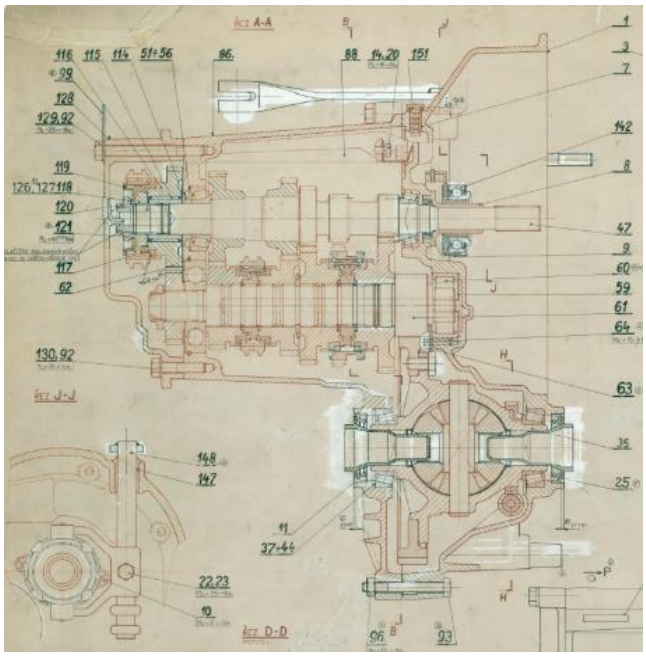


Fig. 11. Š Favorit (1987-1985), transmission - technical drawing

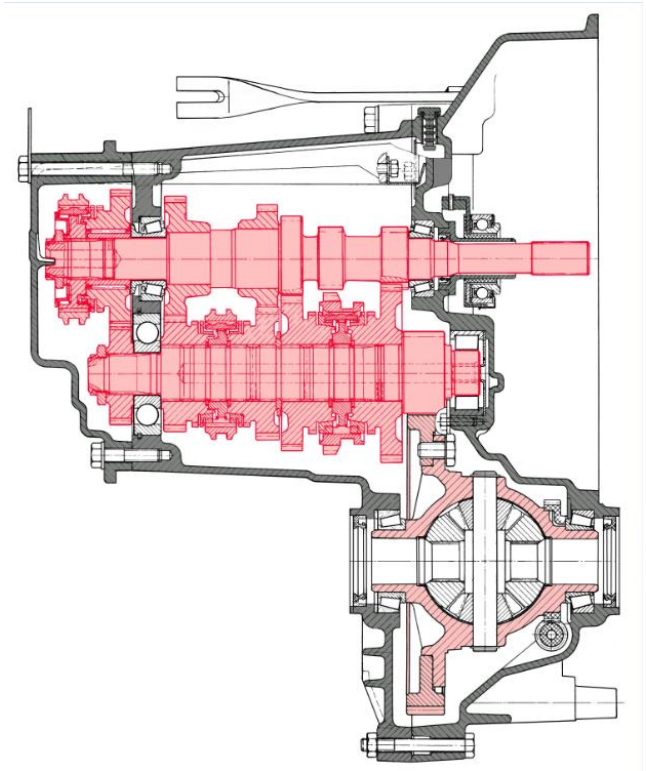


Fig. 12. Š Favorit (1987-1985), transmission - technical drawing, digital modified

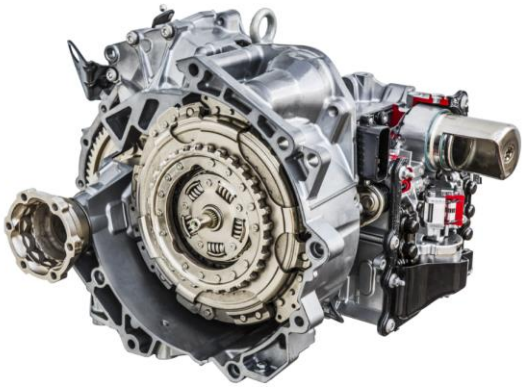


Fig. 13. The new DSG transmission, produced in ŠKODA AUTO plant Vrchlabi

The transmission is normally located in the block with the motor, together with the clutch.

There are also cars that have

1. The motor in front powered rear axle (the classic concept used for Škoda vehicles until 1964, partially until 1973), Fig 1.
2. Motor in rear and powered rear axle (1964-1990), Fig. 5
3. Motor in front with front powered axle (from 1987), Fig. 9-10

An exception is a system in which a transmission is used together with the transfer case at the rear axle, for example with some models of the Popular series in the 1930s.

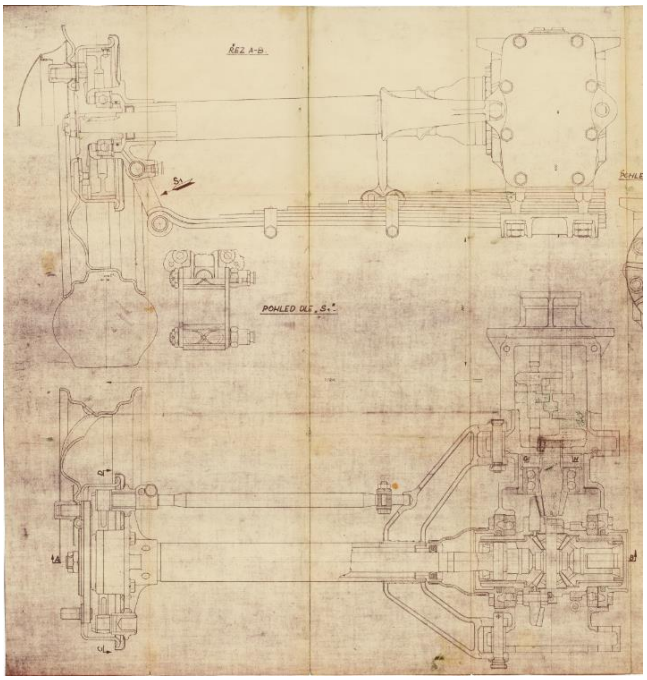


Fig. 14. The gearbox together with differential, rear axle

Acknowledgements

This work was made possible in part by regular activities of the archives of Škoda Auto and in part during the National Cultural Identity project NAKI DG18P02OVV051 Czech Century of Motoring.

References

[1] ŠKODA AUTO Corporate Historical Archives

[1] Type documentation fund

[2] Collection of technical drawings

[2] CEDRYCH René Mario, NACHTMANN Lukáš: ŠKODA - known and unknown cars, second supplemented revised edition, Grada, Prague 2007. ISBN 978-80-247-1719-7

[3] CHLUPÁČ Martin: Laurin & Klement and Škoda history of transmissions. MotoPublic for ŠKODA AUTO 2019. ISBN 978-80-906693-4-5

About Author...

Lukáš NACHTMANN was born in Sokolov in 1973. After secondary school, he studied at the Philosophical Faculty of Charles University in Prague, and later at the Association of Museums and Galleries of the Czech Republic in the field of Museum Studies with a focus on the history of technology. From 2012 to 2019 he was a doctoral student at the Czech Technical University in Prague, in the field of History of Technology (trainers Professor M. Efmertová and Professor Z. Beneš), in which he will resume his studies in the autumn of 2019. Since 1999 he has worked as head of the Škoda Auto Corporate Historical Archives) in Mladá Boleslav.

Looking for existing relations and semantics between test cases and production source code

Matej MADEJA[1]

[1]Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Letná 9, 042 00 Košice, Slovakia

matej.madeja@tuke.sk

Abstract. *Tests are usually an important part of a software project. Their main goal is to ensure product stability and correctness. However, several studies suggest that the tests represent product specification, thus containing information that may not be included in the source code. This paper is focused on finding existing relations between the tests and the source code, and identifying ways developers express semantics in test names. Our results indicate that several tools show the connection of the source code with tests and that the test title often includes use case, which is expressed in the test body implementation in detail. From observed practices, it is obvious that tests can help improve the source code comprehension efficiency.*

Keywords

Github mining, testing practices, test case, source code, program comprehension, unit testing, Android projects.

1. Introduction

Agile methods are very popular [4] in software engineering, mainly due to the incremental delivery of new functionalities that adapt product design to the stakeholders needs. During continuous integration it is important to test new functionality regressively, mostly by manual testing. At the same time, we can see increasing demands for automated tests, which are much more effective than manual tests. Therefore, tests can be considered as a part of quality software and according to Pham et al. [9] their occurrence is expected especially in popular projects.

Demeyer et al. [3] indicates that tests must always be kept up to date with the production code which is tested. Therefore, we can assume that the tests represent current product specification and at the same time we consider it always up-to-date documentation. The developer needs to think about many tasks when writing the source code - implement specified requirements, prepare for possible changes in the future, code optimization, etc. All these facts are processed in developer's head, formed into the so-called *mental model*, and maps it to a specific software artifact (mostly in the form of source code). These facts distract the developer many times and the code may not be written in a form that is easily comprehensible by another developer. On the contrary, during writing tests developer concentrate solely on writing a test from the specification, so it is assumed that the tests will contain information that can help with the understanding of the production code.

In our previous research [7] we encountered the problem of creating consistent materials for students that reflects tests along with source code. Actually, the problem is that we do not know existing relations between tests and source code, which could be used to improve source code comprehension. Therefore, in the following sections we describe the chosen approach, summarize the survey of the relations found between the tests and the production code and observed practices of writing tests in popular Android Github projects.

2. Approach

Because the tests check the source code, it goes without saying that there exists a relationship between these artifacts. Research has not yet verified the possibilities of streamlining the source code comprehension by testing, so it is necessary to observe approaches that are closely related to this topic. In order to describe in detail how and when the tests reflect production code and vice versa, we have analyzed the existing tools commonly used in reverse and software engineering. As a result of this observation, an overview of existing approaches that directly point to the existence of these relations between mentioned software artifacts has been created.

There are number of recommendations on how to write tests effectively, e.g. naming conventions [6] and effort to minimize the occurrence of test smells [11]. Nevertheless, this recommendations do not describe what the developers express in test titles and how it is related to body of the test itself or tested code (UUT). To find test semantics and its relation to the source code semantics, we focused on the real open source Github Android projects. According to Pham et al. [9] tests are expected in popular projects,

because such projects require the reliability of the created code, mostly used by end developers. By manual analysis and looking for naming connections we would like to find out how and whether the developers write tests in real open-source projects. Android platform was chosen due to assumed occurrence of unit, integration and user interface (UI) tests, making it possible to compare with each other. Inspired by article of Kochhar et al. [5], projects were selected from the F-Droid[1] website, a community-maintained software repository for Android. We have ranked projects by Github popularity (number of stars) and a total of 20 most popular projects in February 2019 were used for the analysis. By manual reading of the tests and the production code we tried to comprehend them and look for semantic links of the test title with the test body, and the test title with unit under test (UUT).

3. Observation results

3.1. Overview of relations

By inspection the source code of selected open-source projects and observing the functionality of multiple tools, we have reached 3 types of existing relations between the tests and the production code described below (Fig. 1).

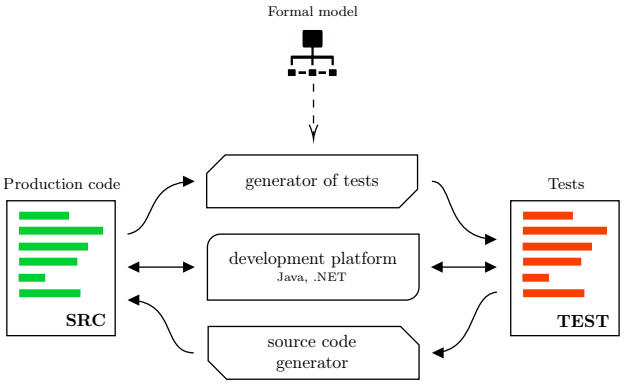


Fig. 1. Existing relations of the source code and the tests. Full arrows indicate the source code impact of the test and vice versa; dotted line - a formal model can be a source for a test generator.

Development platform

Prado et al. [10] mention general awareness that *xUnit* software frameworks are dominant in unit testing (e.g. JUnit for Java). As observed by Atkison et al. [1], tools such as JUnit still require writing tests in the form of programs, which is beneficial to their natural integration with the source code. However, unit tests may not be written in the same language as the production code (e.g. C# code can be tested by a test written in Visual Basic), but they are always written within the same platform (e.g. .NET). This means that links are created between the code and language details, test strategy, test data, and ultimately test results. It is possible to claim can that there exist relations and dependencies between the code and the tests.

Depending on the abstraction of the test method or test type, the connection between the tests and the source code changes. Unit tests are most dependent on the source code because they test the code in detail (e.g. a praticular method), therefore, the test use multiple source code artifacts, such as constants, partial results, etc. User interface (UI) tests have the weakest connection with the production code. Testing code references the production code at most at element identifiers level, for example, to perform a gesture over a particular element. However, even UI tests can refer to the source code at all, because searching a specific element can be based on the displayed content, such as text placed in the element.

Test case generators

According to [2] test code generators can be divided into two main groups: model-based and code-based generators. The goal of these generators is to simplify test creation which is time-consuming. Model-based generators combine the mental model of the tester and the formal model that is created directly from the specification by a developer or a skateholder. As mentioned earlier, the goal of the tests is to verify the program's functionality against specification, so it is more accurate technique than code-based generators, described below. It can create tests without conflicts, ambiguities, structural or behavioral errors because the generation is done according to a specific algorithm that follows the exact meaning of the formal model.

Code-based generators do not use the developer's mental model at all, nor are they dependant on requirements. The production code, which is considered error-free, is used as a data source for generator. UUT should be tested based on the specification and because error-free source code is very difficult to achieve, there is an increased incidence of errors in the tests themselves. Inputs for the tests generation are selected from a text or binary code according to a specific criteria, e.g. in order to achieve maximum test coverage. Both types of generators show that if the source code is complete and absolutely error-free, it is possible to automatically generate a test suite, meaning that the source code contains all the information about tests.

Source code generators

Niznar and Pałka in the recent article [8] describe a method of generating source code from TDD tests. The production code is generated based on data from Abstract Syn-

[1] https://f-droid.org/

tax Tree (AST) enriched with data from static code analysis. However, such test generation is semi-automatic and mostly require manual intervention by the developer because it is not always possible to detect the test semantics. For example, if a class has two parameters whose type cannot be clearly identified, the attribute is generated twice. In ideal conditions, when the test suite is absolutely complete (all test combinations covering the entire specification), it is possible to generate full and exact source code from the tests. These generators confirm the fact that the tests contain information that can describe the tested source code, e.g. by enriching the code with annotations or comments. The biggest problem is the correct recognition of test semantics.

3.2. Testing practices in open-source projects

By reading and searching for the semantics of tests in 20 popular Github Android projects, the following facts have been observed. We divided them into smaller parts to make the context more transparent.

Semantics used in unit test titles

Android projects very often use the name of UUT in its unit test title, as naming conventions for unit tests recommend. However, it is not the rule. There have been cases where the UUT name was not completely included, for example, UUT named `reservation()` has been tested by `reserve()` test. It has also been found that developers do not always use the word `test` in the test title as shown in the example above. Because fixture or helper methods can also be found in the test class, the word `test` in the title can be very helpful in identifying the test method. Since all analysed projects used the JUnit testing framework, which annotates each test with `@Test`, it was solidly possible to accurately identify the test method.

Another way of test naming was using synonyms of words used in UUT, e.g. the `processForm` method was tested by a test titled as `testSubmit`. In this case, it may not be possible to automatically detect the UUT from the title. If the UUT has been tested by several different tests, the test method title contained use case that corresponded to the test body execution. Such a test method usually had the following structure:

```
test<UUTname>_<stateBeforeTest>_<expectedResult>
```

Thanks to this type of test naming it is possible to assume that in most cases we can find also state before test start and expected results. Despite the Java camel-case conventions (e.g. `myMethodName`) commonly used by Android, developers used the '_' character (underscore) to separate meaning groups.

Differences between unit and other test types

We also noticed differences between naming of unit test, integration and UI tests. In integration tests, there were tests that had the same format as unit tests template mentioned above. This type of naming occurred when a test tested a small functionality in the context of a real device. Other type, which has been not seen in unit test, was test title containing only the use case of the test. We also found use cases in unit tests, but in integration tests have been described in more detail, such as `registerUser_whenNotCreatedObject_afterListenerRegistered()`.

UI tests contained the highest quality description of use case from all tests types in their titles, because unit test actions were no longer concerned with the internal structure of the project, but with the use of the application by end users. Therefore, the names contained specific actions expected of the end user in the production environment or what the user would expect from the application, e.g. `mapFragmentShouldBeHidden_With1GpsPoint()`.

Connection between title and body

The words in the test title were very similar to the names of the called and tested method. Most of the words from the test title were also found in its body and partly in the production code which has been tested. This suggests that developers write source code in a way that in which artifacts use similar vocabulary. On this basis, it is possible to identify the UUT more accurately and to look for connections in other directly unrelated parts of the code. Words included in test titles were most often a part of: variables, called methods, or comments.

Usage of statements

Best practices of writing tests claim that only one `assert` should be called in a test. The increased occurrence of test smells in the code can make the program comprehension more difficult. Most tests in our observation included more than one `assert` method call, making it very difficult for us to identify the UUT. In many cases, multiple UUTs or functionalities were tested in one test. Since this fact can have a significant impact on the automatic mining of information from the tests, in the future it will be necessary to minimize this smell or suggest a way to unequivocally identify the UUT despite it.

Occasional use of the `for` loop in the tests has been also related to this fact. In these loops `assert` calls was very common and in some cases more than 20 iterations have been executed. Rarely, we have seen the use of control flow statements, such as `if-else`.

4. Threads to validity

Conducted survey probably does not include all the tools and methods that indicate the connection of the source code and tests. Focusing our research on improving the program comprehension could influenced the perspective of looking for relations between observed artifacts and analyzing existing approaches.

The analyzed projects do not represent tests in all world projects and describe only the status of tests in popular open-source Android projects on Github. The projects have been implemented by different teams that used their own writing conventions and they could be influenced by the similar Github community. Further, these results do not include proprietary projects in which developers have different motivation to create a quality code.

The source code comprehension and creation establish relations between the test and the UUT was dependent on the observer's mental model. To make the results more general, multiple observers should participate. The results presented in this paper are not quantified because the main objective of this observation was to understand how developers write tests in real projects, and whether there are relations that can be quantified and used in the future. Quantification of found relationships will be accomplished in the future.

5. Conclusion and future work

This paper describes the existing relations between tests and production source code that can be useful for reducing effort during source code comprehension tasks in the future. Overall, we found three types of relations: 1) using development platform, 2) test case generators and 3) source code generators. It was found that the largest connection to the source code have unit tests, i.e. they use directly the source code components, e.g. constants, methods, etc. Conversely, the smallest connections have UI tests, where a relation is created at most through UI element identifiers, or the relation between the test and the UUT is created only using printed text on the screen, losing the relation completely.

Test title often inludes the name of UUT or a part of it's name. If a synonym appears in the test title, it may not be possible to automatically identify the UUT or it should be possible with artificial intelligence. Unit tests titles used a stable template by which the test functionality could be predicted. In addition, the words used in the test title were also found in the test body and, of a lower occurrence, in the UUT body. Loops are often used in tests that negatively affect the source code comprehension and often make it impossible to accurately identify the UUT.

From the observed real-world test writing practices and the discovered connections, we can assume that the tests actually contain information to enrich the source code. Using information from the tests (such as semantics and structure) it is possible to enrich the production code and help developers to comprehend the source code more effectively and facilitate product maintenance.

Given obtained results we plan to perform frequency analysis of words in tests' titles and corresponding UUTs on the Github platform in the future with the detailed quantification of the observed practices. At the same time, we would like to focus on quantification of test smells, which negatively influence the comprehension of test semantics, or even make them impossible to comprehend. Subsequently, we plan to enrich the source code, it's documentation, or facilitate the comprehension of the source code using the support IDE tools.

Acknowledgements

Research described in the paper was supervised by Assoc. Prof. Ing. Jaroslav Porubän PhD., DCI FEI TU in Košice. This work was supported by project VEGA No. 1/0762/19: Interactive pattern-driven language development.

References

[1] ATKINSON, C., BARTH, F., AND BRENNER, D. Software testing using test sheets. In *2010 Third International Conference on Software Testing, Verification, and Validation Workshops* (2010), IEEE, pp. 454–459.

[2] CSEPPENTŐ, L., AND MICSKEI, Z. Evaluating code-based test input generator tools. *Software Testing, Verification and Reliability 27* (2017), 1–24.

[3] DEMEYER, S., DUCASSE, S., AND NIERSTRASZ, O. *Object-oriented reengineering patterns*. Elsevier, 2002.

[4] HAMED, A. M. M., AND ABUSHAMA, H. Popular agile approaches in software development: Review and analysis. In *2013 INTERNATIONAL CONFERENCE ON COMPUTING, ELECTRICAL AND ELECTRONIC ENGINEERING (ICCEEE)* (Aug 2013), pp. 160–166.

[5] KOCHHAR, P. S., THUNG, F., NAGAPPAN, N., ZIMMERMANN, T., AND LO, D. Understanding the test automation culture of app developers. In *2015 IEEE 8th International Conference on Software Testing, Verification and Validation (ICST)* (April 2015), pp. 1–10.

[6] LINK, J. *Unit testing in Java: how tests drive the code*. Elsevier, 2003.

[7] MADEJA, M., AND PORUBÄN, J. Automated testing environment and assessment of assignments for android mooc. *Open Computer Science 8*, 1 (2018), 80–92.

[8] NYZNAR, M., AND PAŁKA, D. Generating source code templates on the basis of unit tests. In *Information Systems Architecture and Technology: Proceedings of 37th International Conference on Information Systems Architecture and Technology – ISAT 2016 – Part II* (Cham, 2017), A. Grzech, J. Świątek, Z. Wilimowska, and L. Borzemski, Eds., Springer International Publishing, pp. 213–223.

[9] PHAM, R., SINGER, L., LISKIN, O., FIGUEIRA FILHO, F., AND SCHNEIDER, K. Creating a shared understanding of testing culture on a social coding site. In *Proceedings of the 2013 International Conference on Software Engineering* (2013), IEEE Press, pp. 112–121.

[10] PRADO, M. P., VERBEEK, E., STOREY, M., AND VINCENZI, A. M. R. Wap: Cognitive aspects in unit testing: The hunting game and the hunter's perspective. In *2015 IEEE 26th International*

Symposium on Software Reliability Engineering (ISSRE) (Nov 2015), pp. 387–392.

[11] VAN DEURSEN, A., MOONEN, L., VAN DEN BERGH, A., AND KOK, G. Refactoring test code. In *Proceedings of the 2nd international conference on extreme programming and flexible processes in software engineering (XP2001)* (2001), pp. 92–95.

About Authors...

Matej MADEJA was born in 1992 in Kežmarok, Slovakia. In 2017 he graduated (MSc) at the Department of Computers and Informatics of the Faculty of Electrical Engineering and Informatics at Technical University of Košice. He defended his master's thesis in the field of Informatics. Currently, he is a PhD student in the same department. His research is focused on improvement of program comprehension efficiency, source code testing techniques and environments, and teaching of programming.

Approximate search in genomic data

Petr RYŠAVÝ[1]

[1]Dept. of Computer Science, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

petr.rysavy@fel.cvut.cz

Abstract. *Genomic data are becoming one of the most bulk data in the world. Therefore, there is a need for efficient manipulation and mining of those data. However, with the advances in technology, the researchers often publish sequencing data in the raw form of reads. In this paper, we evaluate the effects of replacing a sequence with read data on a search under the Levenshtein distance. Namely, we provide several experiments that show how well the search works for read data under various conditions.*

Keywords

edit distance, sequence overlap, genetic sequences

1. Introduction

Paper [7] predicts that in a few years genomic data will become the most bulk data in the world, surpassing the amounts of astronomical or stream data. Currently, the sequencing technology is not able to read a whole genome of a single organism at once. Instead, short subsequences called *reads* can be sequenced. Each read is an approximate substring of the true DNA sequence of the organism; however, we do not know where is the location of the read in the DNA sequence.

To obtain the original sequence, the sequencing is done so that each position is covered by multiple reads. The average count of such reads is termed *coverage*, and we will denote it α. Suffix-prefix overlaps then guide the *assembly* process that is capable of the original sequence reconstruction. With the advances of sequencing technologies, more, and more sequences are provided only as raw reads or only first steps of assembly are done.

Therefore, there is a need for algorithms to handle such unprocessed data. In our previous work [4, 5], we proposed a method to estimate the similarity of two sequences given only their read sets. The measure is based on the Monge-Elkan distance [3]. Denote R_A the set of reads that were sequenced from original sequence A. Denote dist the Levenshtein distance [2] between two sequences. Assume that all reads have the same length l. Then the Monge-Elkan distance between two read-sets is defined as

$$\underset{\mathsf{ME}}{\mathrm{Dist}}(R_A, R_B) = \frac{1}{|R_A|} \sum_{i \in R_A} \min_{j \in R_B} \mathrm{dist}(i, j). \qquad (1)$$

With modifications (such as rescaling) presented in [4, 5], this method is simple, yet good approximation of $\mathrm{dist}(A, B)$.

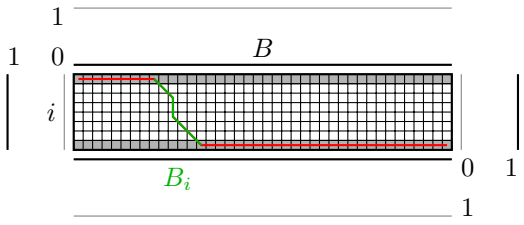


Fig. 1. Modification of the Wagner-Fischer algorithm for search. In sequence B, the leading and trailing gaps are not penalized. The closest part to read i (denoted B_i) is marked with the green line.

2. Contribution

Our goal is to come in the future with theoretical guarantees that explain why (1) is a good approximation of the Levenshtein distance. In this paper, we would like to present an experimental evaluation of the minimum search $\min_{j \in R_B} \mathrm{dist}(i, j)$. For a particular read i, we search for the most similar read in R_B. Assume that an oracle provides us with sequence B. Then finding the minimum $\min_{j \in R_B} \mathrm{dist}(i, j)$ can be viewed as an approximate search for i in sequence B. That means to find a substring of B that has minimum distance to i.

The search can be done by a modification of the Wagner-Fischer algorithm [8]. The penalty for leading or trailing margin gaps in B is replaced with a zero penalty. As a result, the algorithm finds a substring of B that has the lowest distance from the query i. Denote this substring B_i. Notation and gap costs are illustrated in Figure 1. This modification to the Wagner-Fischer algorithm was presented by Sellers in [6].

Applying the same modification to $\mathrm{dist}(i, j)$ inside (1) leads effectively to a sampling of sequence B as anything

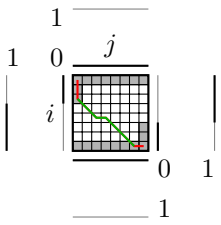


Fig. 2. In the case of read data, we search for i only in j, which is a randomly selected substring of B. Because the length of alignment of B_i and i may be longer than l, we do not penalize first t leading or trailing gaps in i. In the figure, $t = 2$.

besides B_i is not relevant to the final result. Two cases may arise.

- If $|B_i| \leq l$, then one of the reads j may cover B_i and as a result $\mathrm{dist}(i, j) = \mathrm{dist}(i, B_i)$. If no j covers B_i, then the minimum is larger than $\mathrm{dist}(i, B_i)$.
- If $|B_i| > l$, then there exists no read j that covers B_i. However, some reads j may still cover parts of B_i. Therefore, to account for this, we may allow up to t leading or trailing margin gaps of j with zero cost in alignment calculating $\mathrm{dist}(i, j)$. This situation is illustrated in Figure 2.

In this paper, we present an experimental evaluation of the influence of coverage α, the similarity of sequences A and B, and parameter t on difference $\min_{j \in R_B} \mathrm{dist}(i, j) - \mathrm{dist}(i, B_i)$. For notation brevity, denote this difference $\Delta(i)$.

3. Experimental Results

We used ten pairs of sequences to account for different settings of the data. Six pairs included real-world sequences; the remaining four were random sequences (generated under the uniform i.i.d. assumption using a tool on `http://www.faculty.ucr.edu/~mmaduro/random.htm`). Two real-world pairs included similar virus DNA sequences, in the first case they were short, in the second long. Then for both dissimilar sequences, and random sequences we considered all four possible combinations of short and long sequences A and B. The summary of used datasets is in Table 1.

For each pair of sequences, we sampled i.i.d. read-set without replacement for coverage values $\alpha \in \{1, 2, 3, 4, 5\}$. Then we evaluated $\Delta(i)$ for $t \in \{0, 1, \ldots, 20\}$. We set $l = 100$, a common read length provided by the current sequencing technology.

The experimental results are shown in Figures 3, 4, 5. In the case of real-world sequences, the minimum is approximated better for dissimilar sequences. This may be explained by the fact that the minimum is higher and, therefore, a random match is closer to the minimum. There is not much influence on results depending on the choice of whether we search for reads coming from the short sequence in a long sequence or vice versa. For random sequences, the distribution is similar to real-world sequences.

As expected, the average of $\Delta(i)$ decreases with coverage of the data. For each unit increase of t, $\Delta(i)$ is smaller by approx. $\frac{1}{2}$. The effect of t is as expected - the error gets lower as we enlarge t. $\mathrm{dist}(i, B_i)$ is fixed, we, therefore, make the minimum in (1) smaller. The results, however, show that the expected value of $\Delta(i)$ is data dependent.

4. Conclusion

In the paper, we have seen that in 8 out of 10 test cases, the minimum in 1 approximates well $\mathrm{dist}(i, B_i)$. However, in two cases the observed difference was larger. This corresponds to the case of similar real-world sequences and mapping of short random sequence to long random sequence. The distribution of the error $\Delta(i)$ is, however, dependent on data and requires some additional study.

Acknowledgements

The research described in the paper was supervised by Prof. F. Železný, FEE CTU in Prague. This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS17/189/OHK3/3T/13. Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042), is greatly appreciated.

References

[1] Rasko Leinonen, Ruth Akhtar, Ewan Birney, Lawrence Bower, Ana Cerdeno-Tárraga et al. The European Nucleotide Archive. *Nucleic Acids Research*, 39(suppl_1):D28–D31, 2011. doi: 10.1093/nar/gkq967.

[2] Vladimir I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics doklady*, 10(8):707, 1966.

[3] Alvaro E. Monge and Charles P. Elkan. The field matching problem: Algorithms and applications. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, pages 267–270. AAAI Press, 1996.

[4] Petr Ryšavý and Filip Železný. Estimating sequence similarity from read sets for clustering sequencing data. In Henrik Boström, Arno Knobbe, Carlos Soares and Panagiotis Papapetrou, editors, *Advances in Intelligent Data Analysis XV: 15th International Symposium, IDA 2016*, pages 204–214. Springer International Publishing, 2016. doi: 10.1007/978-3-319-46349-0_18.

[5] Petr Ryšavý and Filip Železný. Estimating sequence similarity from read sets for clustering next-generation sequencing data. *Data Mining and Knowledge Discovery*, 33(1):1–23, Jan 2019. doi: 10.1007/s10618-018-0584-8.

[6] Peter H Sellers. The theory and computation of evolutionary distances: Pattern recognition. *Journal of Algorithms*, 1(4):359 – 373, 1980. doi: https://doi.org/10.1016/0196-6774(80)90016-4.

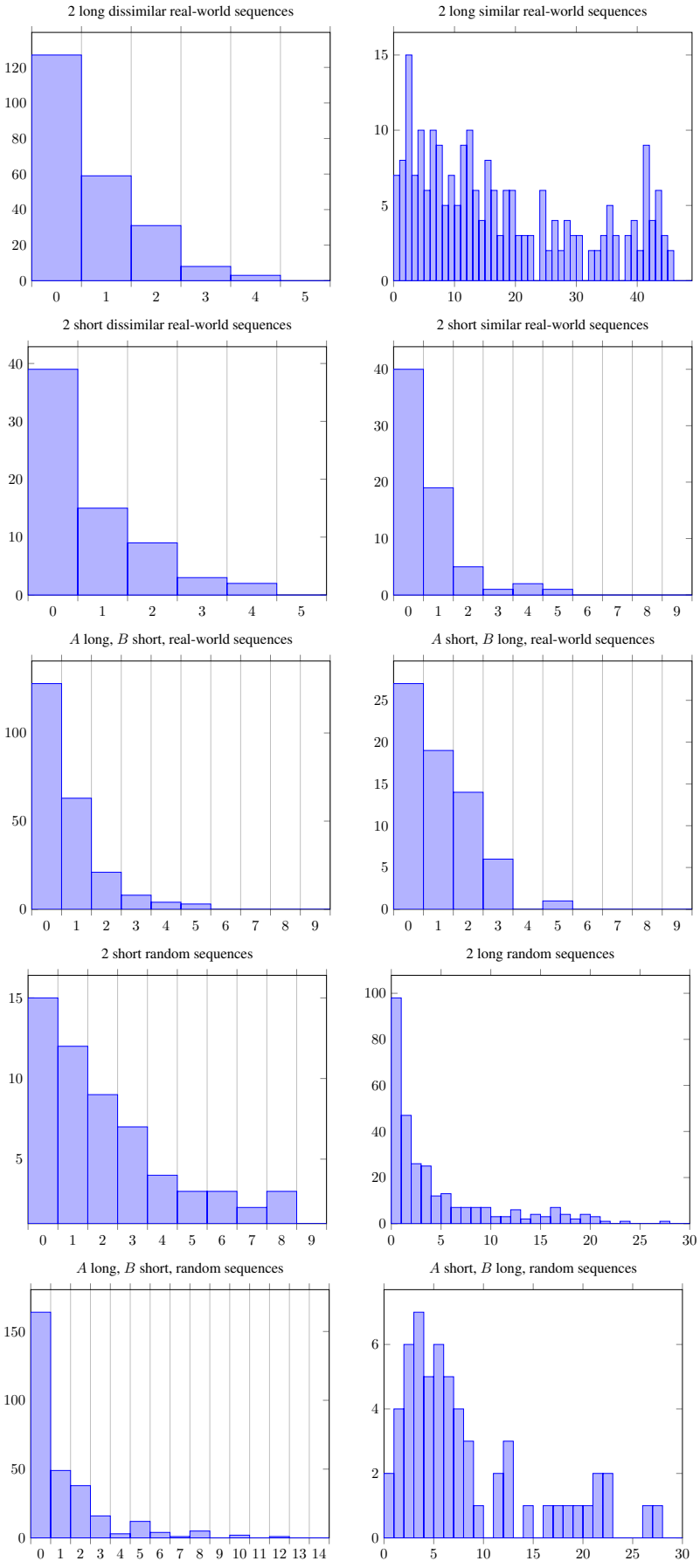


Fig. 3. Histogram of error $\Delta(i)$ for $\alpha = 3$ and $t = 0$.

Dataset description	A	B	$\|A\|$	$\|B\|$
2 long dissimilar real-world sequences	DQ812094	AB447435	7702	7511
2 long similar real-world sequences	DQ812094	DQ219396	7702	7707
2 short dissimilar real-world sequences	FJ966079	EF015778	2280	2687
2 short similar real-world sequences	FJ966079	FJ966080	2280	2274
A long, B short, real-world sequences	DQ812094	FJ966080	7702	2280
A short, B long, real-world sequences	FJ966079	DQ219396	2280	7707
2 short random sequences	random	random	2000	2000
2 long random sequences	random	random	10000	10000
A long, B short, random sequences	random	random	10000	2000
A short, B long, random sequences	random	random	2000	10000

Tab. 1. Description of the datasets used. The code under the real-world datasets refers to ENA [1] accession, by which it is possible to find the sequence online.

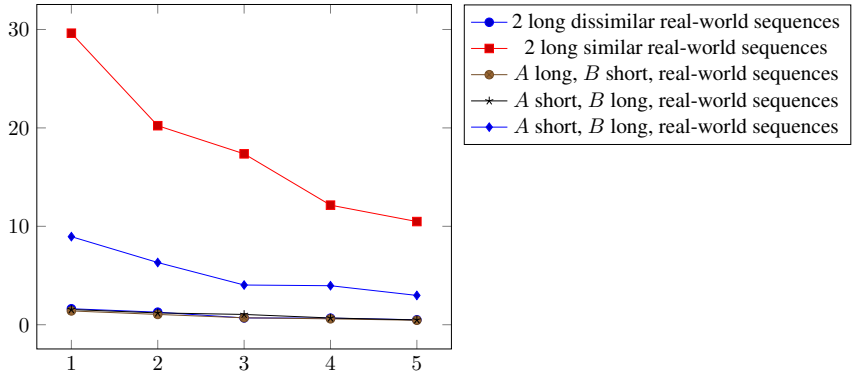


Fig. 4. Dependence of error $\Delta(i)$ on coverage α for $t = 0$.

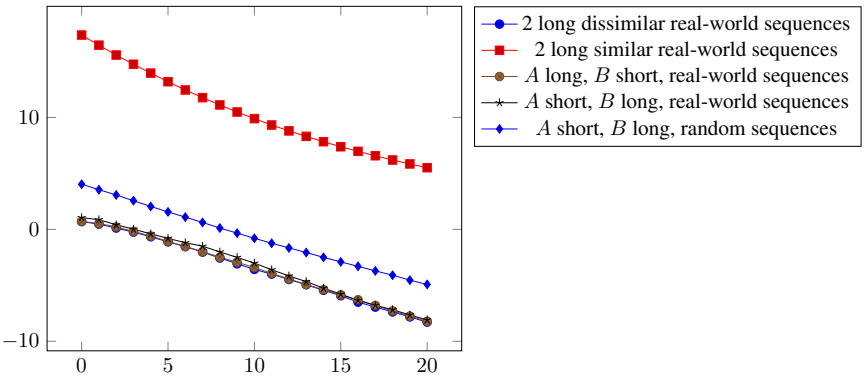


Fig. 5. Dependence of error $\Delta(i)$ on parameter t for $\alpha = 3$.

[7] Zachary D. Stephens, Skylar Y. Lee, Faraz Faghri, Roy H. Campbell, Chengxiang Zhai et al. Big data: Astronomical or genomical? *PLOS Biology*, 13(7):1–11, 07 2015. doi: 10.1371/journal.pbio.1002195.

[8] Robert A. Wagner and Michael J. Fischer. The string-to-string correction problem. *Journal of the Association for Computing Machinery*, 21 (1):168–173, January 1974. doi: 10.1145/321796.321811.

Hybrid Code Networks using a convolutional neural network as an input layer achieves higher turn accuracy

Petr Marek

Dept. of Cybernetics, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

marekp17@fel.cvut.cz

Abstract. *The dialogue management is a task of conversational artificial intelligence. The goal of the dialogue manager is to select the appropriate response to the conversational partner conditioned by the input message and recent dialogue state. Hybrid Code Networks is one of the models of dialogue managers, which uses an average of word embeddings and bag-of-words as input features. We perform experiments on Dialogue bAbI Task 6 and Alquist Conversational Dataset. The experiments show that the convolutional neural network used as an input layer of the Hybrid Code Network improves the model's turn accuracy.*

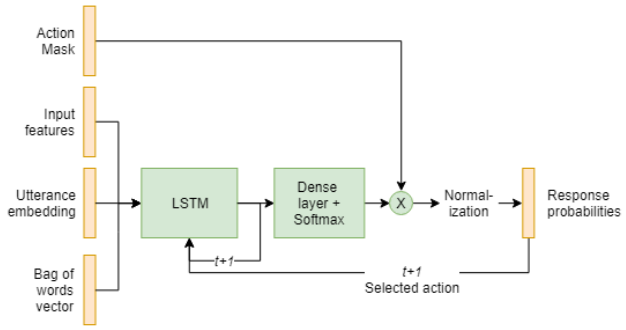


Fig. 1. Schema of baseline Hybrid Code Networks

Keywords

Conversational artificial intelligence, dialogue management, Hybrid Code Networks, convolutional neural network

1. Introduction

There was a significant spread of personal assistants and chatbots in recent years. Users have greater demands on their capabilities, as they become more mainstream and popular. The demand drives the research of better technologies of conversational artificial intelligence. One of the technologies is dialogue managers. The role of the dialogue manager is to select the most appropriate response based on the recent user's message and state of the dialogue.

Hybrid Code Networks is a type of dialogue manager, which combines a recurrent neural network with the domain-specific rules. Rules allow the model to learn with a considerably reduced amount of training examples.

The baseline Hybrid Code Networks uses an average of word embeddings and bag-of-words of user's message as input features. These features do not capture the order of the words in the message. We propose two new architectures inspired by Hybrid Code Networks, which can capture the order of words. The first architecture uses a convolutional neural network and the second architecture uses a recurrent neural network as input layers.

2. Related work

There are various approaches to dialogue management. Dialogue managers can be divided into retrieval and generative dialogue managers, and into rule-based, end-to-end and hybrid-based dialogue managers.

The rule-based dialogue managers produce a response based on the set of rules [7]. The rules detect the presence of patterns in the user's input messages and the state of the dialogue. The rule-based systems work well in constrained domains, but they become more inaccurate and more challenging to implement as the complexity of the domain grows.

Dialogue managers working with text similarity measures the similarity between the user's input message and set of reference sentences. The most similar sentence out of the set of reference sentences determines the next action of the dialogue manager. The similarity is measured between vector representations of the message and sentences [7]. Vector representations can be created by TF-IDF or average of word embeddings like word2vec [8], GloVe [9] or fastText [10]. The similarity function can be cosine similarity or function learned by supervised learning [11].

Seq2seq dialogue managers learn a mapping between input message and output response [12]. It consists of encoder and decoder recurrent neural networks. The encoder maps input message into a fixed size vector representation. The decoder generates response out of the fixed size vector representation. The seq2seq models were initially used for

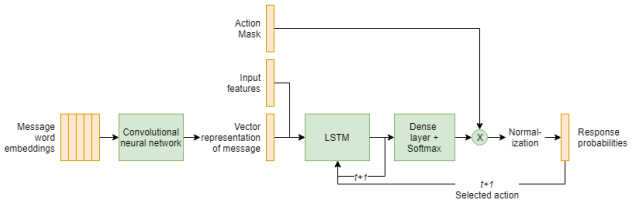


Fig. 2. Schema of Hybrid Code Networks with the convolutional input layer

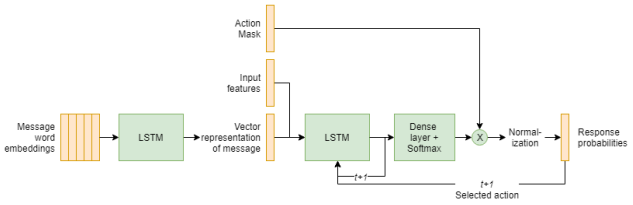


Fig. 3. Schema of Hybrid Code Networks with the recurrent input layer

machine translation. However, we can use them as generative dialogue managers [13].

Memory networks are a class of dialogue managers, which combines inference component with a long term memory. Memory network can read from and write to the memory. The model learns how to operate the memory efficiently. A memory network consists of four components, which can be trained separately [14] or end-to-end [15].

3. Model Architectures

We propose two new architectures of Hybrid Code Networks. The baseline model of Hybrid Code Networks is displayed in Figure 1. It uses an average of word embeddings (utterance embedding) and bag-of-words as input features.

The first proposed architecture uses a convolutional neural network to create input features instead of an average of word embeddings and bag-of-words. The architecture is displayed in Figure 2. The architecture of the convolutional neural network is described in detail in [6]. The only change we did to the convolutional neural network was removing the classification layer. The convolutional neural network is trained jointly with the rest of the model of the dialogue manager. The second architecture uses a recurrent neural network to create input features. We use LSTM recurrent cell [16]. The architecture is displayed in Figure 3. Both proposed model architectures take input message as a list of word embedding vectors.

4. Datasets

We tested the proposed architectures on Dialogue bAbI Task 6 [7] and Alquist Conversational Dataset.

The Dialogue bAbI Task 6 dataset consists of real human-bot dialogues and knowledge base from the restauration reservation domain. The dataset consists of 3,200 training dialogues, 400 validation dialogues, and 400 testing dialogues. The task is to select one of 56 possible responses.

The Alquist Conversational Dataset consists of conversations between users and the socialbot Alquist [17]. It is a non-public dataset. The dataset consists of 37,000 dialogues about books. There are 340,000 message-response pairs in total. The average length of dialogues is 9.11 pairs, the median is 7 pairs, and there are 23,000 unique responses. The responses are clustered into 30 semantically similar classes. The goal is to select one of 30 response classes. The dataset is noisy and hard to learn because it contains voice recognition errors and part of the messages come from uncooperative users. Messages from uncooperative users are hard to interpret or out of the domain of books.

5. Experiments

5.1. Experiments on Dialogue bAbI Task 6

We tested the proposed architectures on Dialogue bAbI task 6 two times, each time with different set of embedding vectors. We used the pretrained 300-dimensional word2vec embedding vectors trained on News[1] and 300-dimensional fastText embedding vectors pretrained on the training set of Dialogue bAbI Task 6 dataset. The fastText embeddings were trained for 100 epochs. We fixed the values of both types of embeddings during training of whole models.

We performed the hyperparameter optimization using Bayesian optimization on the validation part of the dataset. We used the implementation from Scikit-Optimize library[2]. We performed 30 rounds of training with a different set of hyperparameters for each model. Each training was 30 epochs long. We evaluated the performance of the model after each epoch and saved the weights of the model achieving the highest turn accuracy on the validation set for each set of hyperparameters. We selected the hyperparameters achieving the highest turn accuracy. The values of selected hyperparameters are presented in table 1.

We trained each model with the hyperparameters, which achieved the highest validation accuracy. We used 12 training epochs. We measured the turn accuracy on validation set after each training epoch and saved the weights achieving the highest validation turn accuracy. We used these weights for testing. Results are presented in section 6.

[1] GoogleNews-vectors-negative300.bin
[2] https://scikit-optimize.github.io/

Parameter/Model	fastText	fastText+CNN	fastText+RNN	word2vec	word2vec+CNN	word2vec+RNN
LSTM size	55	245	505	85	109	219
Input LSTM size	-	-	199	-	-	312
Convolutional filters	-	21	-	-	6	-
LSTM dropout	0.85	0.80	0.94	0.92	0.79	0.74
Input LSTM dropout	-	-	0.97	-	-	0.91
Convolutional dropout	-	0.72	0	-	0.84	-
Fully connected dropout	0.82	0.79	0.76	0.59	0.93	0.98
Learning rate	0.008	0.0001	0.0003	0.001	0.005	0.00005
Activation function	relu	relu	relu	tanh	tanh	relu
Input activation function	-	-	tanh	-	-	tanh
Adam epsilon	1E-8	1E-8	1E-8	1E-8	0.1	1E-8
Adam beta1	0.9	0.5	0.5	0.5	0.5	0.9
Turn accuracy	69.4%	71.5%	68.0%	71.3%	70.4%	65.5%

Tab. 1. Model's hyperparameters achieving the highest turn accuracy on teh validation set of Dialogue bAbI Task 6

5.2. Experiments on Alquist Conversational Dataset

We measured the performance of all models on the Alquist Conversational Dataset. We used the word2vec embedding vectors pretrained on News and fastText embedding vectors trained for 100 epochs on the training set of Alquist Conversational Dataset. We did not perform hyperparameter optimization due to a long time of training. We used hyperparameters presented in section 5.1 for each model instead.

We trained models for 12 epochs. We measured the turn accuracy after each epoch and saved the weights which achieved the highest turn accuracy. We measured turn and dialogue accuracies on the testing set using saved weights. Results are presented in section 6.

6. Results

We present the accuracy of described models in table 2. Baseline Hybrid Code Networks outperforms other models on the Dialogue bAbI task 6 in both turn and dialogue accuracy [1]. The Hybrid Code Networks using a convolutional neural network as input layer and fastText embeddings improves the turn accuracy of the baseline Hybrid Code Networks by 3.3% on the Dialogue bAbI task 6 dataset. The dialogue accuracy of this model is smaller by 1.4%. However, we were not able to reproduce the dialogue accuracy described in [1] with our baseline model (labeled as word2vec).

Hybrid Code Network model achieves the highest turn accuracy of 92.6% on the Alquist dataset if the model uses a convolutional neural network as an input layer and word2vec embeddings. This model achieves the second highest dialogue accuracy.

The results show that convolutional neural network as an input layer in Hybrid Code Networks achieves higher turn accuracy than baseline Hybrid Code Networks and Hybrid Code Network using a recurrent neural network as an input layer.

7. Conclusion

We proposed two new architectures of Hybrid Code Networks. They use convolutional and recurrent neural networks as input layers, instead of the average of word embeddings and bag-of-words. We tested the turn and dialogue accuracy of the proposed architectures on the Dialogue bAbI task 6 and Alquist Conversational Dataset.

Results show that Hybrid Code Networks using a convolutional neural network as input layer improves the turn accuracy. The model using convolutional neural network outperforms the baseline model of Hybrid Code Networks on the Dialogue bAbI task 6 and Alquist Conversational Dataset.

Acknowledgements

The research described in the paper was supervised by Ing. J. Šedivý, CSc. CIIRC CTU in Prague and Ing. Václav Chudáček, Ph.D. CIIRC CTU in Prague, and supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS19/091/OHK3/1T/37.

References

[1] WILLIAMS, Jason D.; ASADI, Kavosh; ZWEIG, Geoffrey. Hybrid code networks: practical and efficient end-to-end dialog control with supervised and reinforcement learning. arXiv preprint arXiv:1702.03274, 2017.

[2] BORDES, Antoine; BOUREAU, Y.-Lan; WESTON, Jason. Learning end-to-end goal-oriented dialog. arXiv preprint arXiv:1605.07683, 2016.

[3] LIU, Fei; PEREZ, Julien. Gated end-to-end memory networks. In: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers. 2017. p. 1-10.

[4] ERIC, Mihail; MANNING, Christopher D. A copy-augmented sequence-to-sequence architecture gives good performance on task-oriented dialogue. arXiv preprint arXiv:1701.04024, 2017.

[5] SEO, Minjoon; HAJISHIRZI, Hannaneh; FARHADI, Ali. Query-regression networks for machine comprehension. arXiv preprint

Model	bAbI6		Alquist	
	Turn Acc.	Dialogue Acc.	Turn Acc.	Dialogue Acc.
Bordes and Weston (2017) [2]	41.1%	0.0%	-	-
Liu and Perez (2016) [3]	48.7%	1.4%	-	-
Eric and Manning (2017) [4]	48.0%	1.5%	-	-
Seo et al. (2016) [5]	51.1%	-	-	-
Williams, Asadi and Zweig (2017) [1]	55.6%	**1.9%**	-	-
fastText	57.6%	0.8%	86.9%	51.7%
fastText+CNN	**58.9%**	0.5%	90.6%	63.0%
fastText+RNN	54.9%	0.3%	80.6%	40.5%
word2vec	57.4%	0.4%	92.2%	**68.0%**
word2vec+CNN	56.3%	0.1%	**92.6%**	67.8%
word2vec+RNN	54.6%	0.1%	83.9%	45.2%

Tab. 2. Testing accuracy of dialogue managers on Dialogue bAbI task 6 and Alquist Conversational Dataset

arXiv:1606.04582, 2016.

[6] KIM, Yoon. Convolutional neural networks for sentence classification. arXiv preprint arXiv:1408.5882, 2014.

[7] BORDES, Antoine; BOUREAU, Y.-Lan; WESTON, Jason. Learning end-to-end goal-oriented dialog. arXiv preprint arXiv:1605.07683, 2016.

[8] MIKOLOV, Tomas, et al. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781, 2013.

[9] PENNINGTON, Jeffrey; SOCHER, Richard; MANNING, Christopher. Glove: Global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). 2014. p. 1532-1543.

[10] BOJANOWSKI, Piotr, et al. Enriching word vectors with subword information. Transactions of the Association for Computational Linguistics, 2017, 5: 135-146.

[11] BAI, Bing, et al. Supervised semantic indexing. In: Proceedings of the 18th ACM conference on Information and knowledge management. ACM, 2009. p. 187-196.

[12] SUTSKEVER, Ilya; VINYALS, Oriol; LE, Quoc V. Sequence to sequence learning with neural networks. In: Advances in neural information processing systems. 2014. p. 3104-3112.

[13] VINYALS, Oriol; LE, Quoc. A neural conversational model. arXiv preprint arXiv:1506.05869, 2015.

[14] WESTON, Jason; CHOPRA, Sumit; BORDES, Antoine. Memory networks. arXiv preprint arXiv:1410.3916, 2014.

[15] SUKHBAATAR, Sainbayar, et al. End-to-end memory networks. In: Advances in neural information processing systems. 2015. p. 2440-2448.

[16] HOCHREITER, Sepp; SCHMIDHUBER, Jrgen. Long short-term memory. Neural computation, 1997, 9.8: 1735-1780.

[17] PICHL, Jan, et al. Alquist 2.0: Alexa Prize Socialbot Based on Sub-Dialogue Models.

About Authors...

Petr MAREK is a Ph.D. student of conversational artificial intelligence at Faculty of Electrical Engineering, CTU. He works on conversational AI Alquist, which was the second prize winner of Amazon Alexa Prize 2017 and 2018. He finished his master degree in 2018 in artificial intelligence at FEE CTU in 2018. The topic of his master thesis was Dialog manager for conversational AI.

A Simulink-Blockset to Control Lego Mindstorms using Raspberry Pi

Isshu ARAKI[1], Benjamin HENTZE[2]

[1]Department of Electrical Engineering and Information Technology, RWTH Aachen University, Templergraben 55, 52062 Aachen, Germany
[2]Medical Information Technology, RWTH Aachen University, Pauwelsstr. 20, 52074 Aachen, Germany

isshu.araki@rwth-aachen.de, hentze@hia.rwth-aachen.de

Abstract. *The LEGO Mindstorms hardware is used at the RWTH Aachen University for educational purposes in undergraduate studies. It allows students to design and operate simple robotic systems. For the improvement of flexibility, the Raspberry Pi can be used as a faster alternative to the LEGO Mindstorms Brick, which in combination with the BrickPi allows us to control both actuators and sensors. In this paper, a MATLAB Simulink-Blockset is presented, which makes it possible to control LEGO Mindstorms using the Raspberry Pi. Interfacing LEGO Mindstorms with the Raspberry Pi based on Simulink might open up novel possibilities for advanced control engineering courses in graduate studies.*

Keywords

Raspberry Pi, MATLAB Simulink, LEGO Mindstorms, Control Engineering.

1. Introduction

Low-cost hardware like the Raspberry Pi or LEGO Mindstorms is used to introduce students to some real research problems. Using LEGO Mindstorms, simple mechatronic systems can be realized easily which describe important control engineering problems, for example the control of an inverted pendulum [1]. For the analysis of complex control systems however, it is desirable to have more flexibility in both software and hardware. A Simulink support package for interacting with the LEGO Mindstorms Brick [2] is available, which makes implementing control algorithms easier. However, one major limitation of using the Brick is that it can only interact with its proprietary hardware, which makes it harder to extend the system with other sensors or actuators.

The Raspberry Pi 3 Model B is therefore chosen as an alternative to the LEGO Mindstorms EV3 Brick, because of its higher clock rate, bigger memory size (see Tab. 1) and hardware extensibility. The usage of external hardware widens the variety of systems to analyze. Additionally, it is possible to interface the LEGO Mindstorms with the Raspberry Pi using a hardware extension called the BrickPi [3], which is produced and distributed by Dexter Industries. Similarly to the LEGO Brick, it features four sensor and four motor ports. Using the BrickPi and the Simulink support package for the Raspberry Pi [4] it is possible to create control algorithms which can run on the Raspberry Pi to control LEGO Mindstorms. In the following, a possible method for this is presented. The created Simulink-Blockset is provided online [5].

The paper is structured as follows: Section 2 presents the methods used for creating the Simulink-Blockset and the analysis of performance limits. In Section 3 the results are discussed and a conclusion is formulated in Section 4.

	Raspberry Pi 3 B	Lego Brick EV3
Clock rate	1.2 GHz	300 MHz
RAM	1 GB	64 MB

Tab. 1. Comparison of hardware properties. The Raspberry Pi is four times faster and has a much larger memory size.

2. Methods

The communication between the Raspberry Pi and the BrickPi is done via the Serial Peripheral Interface (SPI). The Raspberry Pi acts as the master and sends commands to the BrickPi, which responds as the slave by sending data back. A Simulink-Blockset is created, which automates all communication processes. Thus, in order to interact with the LEGO hardware, it is only necessary to drag blocks from the blockset and drop them into the main Simulink program (see Fig. 1).

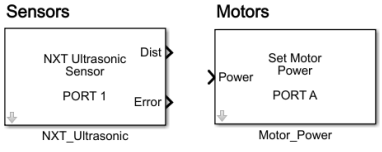


Fig. 1. Example blocks. Left: Block for reading out ultrasonic sensor values. Right: Block for setting motor power.

2.1. Realization of the Blockset

The support package [4] includes the ”SPI Master Transfer” block (see Fig. 2) which enables sending and receiving data bytes. Using this block and the commands which are needed to interact with each individual LEGO hardware, a block for each sensor and actuator is created. The commands are extracted from the C library of the BrickPi which is provided by Dexter Industries. Almost every block consists of an initialization subsystem, which is executed only once at the beginning of the program, and a continuously executed subsystem which retrieves the data (see Fig. 3). Similarly to the blockset which is implemented in the support package [2], masks are added to each block to provide a convenient way for setting parameters.

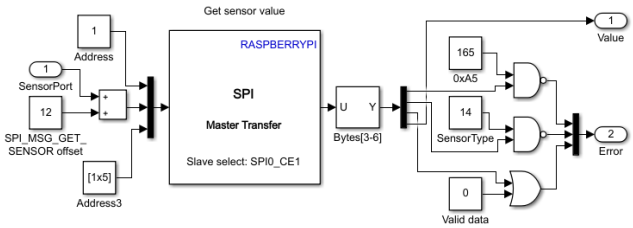


Fig. 2. Example of a communication subsystem. Byte transmission and reception via ”SPI Master Transfer” block.

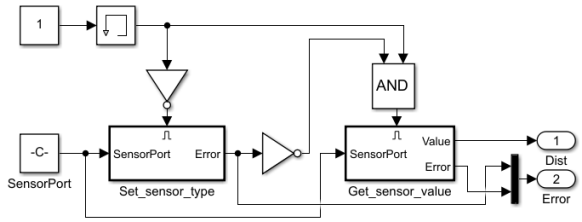


Fig. 3. Internal of a sensor block. Left subsystem is executed once at the beginning for initialization. Right subsystem is continuously executed.

2.2. Analysis of performance limits

For applications in control engineering, it is important to achieve real-time capability, which is why a measurement of maximum possible execution frequency of the blockset is performed. However, because direct measurement of the minimal realizable execution time in Simulink is assumed to be difficult, an approximation is used. The time step size of the Simulink program is decreased until a task overrun is detected at the Raspberry Pi. This is done with every LEGO motor and sensor separately and also with full load, i.e. with four sensors and motors attached.

3. Results and Discussion

The performance analysis shows that the block with the longest execution time needs to be executed with a program step size of at least 1.25 ms in order to avoid task overruns. In case of full load, at least 4 ms are needed (see Tab. 2). Thus, in theory, most control programs can be executed with a frequency of about 250 Hz. However, since the Raspberry Pi is not a real-time system, it does not guarantee this and the given limits should not be seen as tight bounds. As proof-of-concept, the created blocks are used to control a Segway-like robot making it stand in an upright position. The control program runs at 250 Hz.

	Minimal time step
Block	**1.25 ms**
Full load	**4 ms**

Tab. 2. Approximated lower bounds of time step. At full load, a maximum frequency of 250 Hz results.

4. Conclusion

The created Simulink-Blockset enables interaction with LEGO Mindstorms via the Raspberry Pi. Some blocks for certain sensors are not available yet, but the blockset is enough to provide a more flexible tool for advanced control engineering classes in graduate studies.

Acknowledgements

The project described in this paper was supervised by Prof. Dr.-Ing. Dr. med. Dr. h.c. Steffen Leonhardt, Medical Information Technology, RWTH Aachen University.

References

[1] KANADA T., WATANABE Y., GAN C., Robust H2 control for two-wheeled inverted pendulum using LEGO Mindstorms, *2011 Australian Control Conference*, 2011, p. 136 - 141.

[2] MATHWORKS, Simulink Support Package for LEGO MINDSTORMS EV3 Hardware, https://www.mathworks.com/hardware-support/lego-mindstorms-ev3-simulink.html, Accessed: 2019-03-18.

[3] DEXTER INDUSTRIES, BrickPi3 Tutorials & Documentation, https://www.dexterindustries.com/brickpi3-tutorials-documentation, Accessed: 2019-03-18.

[4] MATHWORKS, Raspberry Pi Support from Simulink, https://www.mathworks.com/hardware-support/raspberry-pi-simulink.html, Accessed: 2019-03-18.

[5] ARAKI I., HENTZE B., Simulink for BrickPi, https://git.rwth-aachen.de/medit/simulink-brickpi/, Accessed: 2019-03-22.

About Authors...

Isshu ARAKI is working towards his master's degree in Electrical Engineering, Information Technology and Computer Engineering with a major in ”Systems and Automation” at the RWTH Aachen University.

Datasets Generation for Neural-Symbolic Integration

Martin SVATOŠ

Dept. of Computer Science, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

svatoma1@fel.cvut.cz

Abstract. *Neural-symbolic learning cycle is a machine learning paradigm for learning predictive models by a combination of both artificial neural networks and symbolic rules. There is a vast number of proposed methods for each part of the cycle, yet a proper experimental evaluation, e.g. asymptotic behavior, of the cycle is still missing. One key component for such an evaluation is to have a dataset with expert knowledge to verify whether the cycle produces meaningful results. The aim of this paper is to fill this gap by introducing novel artificial datasets which, in turn, can be used for a proper experimental evaluation of the cycle.*

Keywords

neural-symbolic integration, machine learning, dataset generation

1. Introduction

Our machine learning setting is to classify binary valued vectors, which represent truth values of propositional variables. To solve this task we investigate neural-symbolic integration (NSI) [1] field which combines logic and artificial neural networks (ANNs). As logic and ANNs are very different fields, they possess different advantages and disadvantages. Machine learning models based purely on logic are easy to interpret and encode, but they fail easily on real-world problems containing noise due to the crispness of logic. On the other hand, ANNs-based models can handle noise data but are hard, sometimes even impossible, to interpret due to nonlinear transformations. The core idea of NSI is to combine these two paradigms by two possible means: i) to refine logic programs by encoding them into an ANN and to learn weights to make it more robust to noise, ii) to learn an interpretable logic program from a given ANN. The first part is called *neural phase* since it transfers a logic program to an ANN and refines the network, i.e. refines its weights. The second part is called *extraction phase* since it extracts a logic program from a given ANN. There have been a vast number of methods proposed to solve one of the phases over the last three decades.

Neural-learning cycle is a repetition of these two phases in a loop. The first neural phase of the cycle builds an ANN based on a logic program, i.e. *initial knowledge*, given by a user who should have some expert knowledge in the data. The concept of the cycle was firstly introduced in [5], but up till now, there have been only two experiments in the literature considering one loop from the neural phase to the extraction phase and back [5, 3]. There are still missing answers to the behavior of the cycle with more than just one repetition. These experimental questions are: i) Is the cycle robust to initial knowledge? ii) Does the cycle converge when it learns during multiple iterations? iii) Does the cycle improve symbolic models produced in the extraction phase? iv) Is there any change in the complexity of the symbolic models, e.g. description length, produced in the extraction phase? These questions are still unanswered.

For a proper experimental evaluation of the cycle, we need datasets with initial knowledge, underlying target concept, information about the level of noise in the data, and some score indicating the level of difficulty of the target concept, e.g. description length, since we want datasets of varying difficulty. The initial knowledge can be provided by an expert, however, such an expert can be very costly or does not have to exist at all. While taking into account real-world datasets containing noise, there it is almost impossible to get the initial knowledge. One possible way is to use a standard machine learning methods to learn some rules describing regularities in the data, but that does not tell us what the target concept is. Therefore, we propose a novel way of generating datasets for NSI cycle evaluation. The main advantage of such datasets is that a dataset with target concept, a level of noise, etc., can be generated upon a request. These datasets, in turn, can be used for a proper experimental evaluation of the NSI cycle paradigm and to answer the questions stated above.

2. Related Work

To the best of our knowledge, there are only two artificial datasets that were used for experimental evaluation of NSI methods. The first is *lenses* dataset [2] with four attributes, three classes, and only 24 examples, which is an insufficiently small number for learning weights of an ANN. The second is a family of *monk's* datasets [6] with six attributes, two classes, and 432 examples. Attributes of the monk's dataset correspond to a body and head shapes,

whether a particular monk holds something or smiles, etc. Target concepts based on these attributes were created to generate three separate monk's problem; e.g. one target concept is to classify as positive examples only monks with a bold face holding a sword. Further, noise was introduced to mimic real-world datasets. However, monk's datasets were originally developed for an experimental comparison of symbolic methods, e.g. decision trees, by aiming at their weaknesses, and thus it can be easy to solve the by neural networks.

Describing real-world datasets used in NSI, the most noticeable ones are breast cancer [7] and promoters [4]. However, we lack an expert's initial knowledge; thus they are not much of a use for our problem. Also, the level of noise is not well described in the literature.

As ANNs based method need a large number of examples in order to learn well, it is not the case for datasets used in NSI nor the ones stated above. Even the largest one, breast cancer dataset, has *only* seven hundred of examples. All datasets except the monk's family of datasets lack any kind of description of their difficultness, initial knowledge or target concept.

3. Family of *cnfX-Y* Datasets

In this section, we present a novel family of datasets called *cnfX-Y* which is similar to monk's family. In fact, cnfX-Y can be seen as a generalization of monk's family because the target concept of the latter is just a conjunction of propositional variables. Contrary to that, cnfX-Y datasets possess a target concept of a *CNF*, which is a conjunction of disjunctions. An arbitrary propositional logic formula can be transformed into a CNF using De Morgan's laws. The parameter *Y* expresses the number of propositional variables that can be used, i.e. it is the size of the sample.

We propose a scoring function that will produce a number expressing a difficultness of a dataset. Recall that in a cnfX-Y dataset the underlying formula from which its samples were generated is known. The formula can be easily transformed into a CNF to compute its number of disjunctions, which is our measurement of difficultness of a dataset; this value is denoted as *X*. The CNF should be the minimal one possible in the sense of the number of disjunctions it consists of. The rationale behind this comes from the fact that NSI methods transform a logic program into an ANN by expressing conjunctions and disjunctions by separate nodes. Following this encoding, a CNF can be encoded into a three-layer ANNs; the first layer is the input one, the second (hidden) layer contains only *or* nodes, finally, the third (output) layer consists of only one *and* neuron. Thus, the value of *Y* approximately expresses how large ANNs is needed in order to learn the target concept.

While generating the target concepts, i.e. propositional formulae, we use standard logical connectives as $\wedge$, $\vee$, $\neg$, $\Longrightarrow$, $\oplus$, etc. Further, we put other requirements on the concept space, e.g. maximal and minimal height of syntax tree of a formula. Besides knowing the underlying formula for a dataset, we can control the level of noise, i.e. number of incorrectly labeled data. However, one of the most useful settings during the dataset generation is to control the balance of positive and negative samples, i.e. a dataset where majority class has about 90% of samples can be useless for NSI methods comparison. Besides that, each dataset generated from cnfX-Y family has 2^Y samples, thus sufficient amount of data is available for learning ANNs based methods.

4. Conclusion

This paper presents a novel family of datasets for NSI field and the NSI cycle in particular. Its advantage over current datasets lays in possession of the target concept, easy initial knowledge construction from the concept, information about the level of noise, and a scoring function describing difficultness of a dataset from this family. However, these datasets can be used in an arbitrary classification task in machine learning, e.g. decision tree learning, and are thus not subjected only to the field of NSI.

Acknowledgements

The research described in the paper was supervised by Gustav Šourek, FEE CTU in Prague. This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS17/189/OHK3/3T/13. Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme Projects of Large Research, Development, and Innovations Infrastructures (CESNET LM2015042), is greatly appreciated.

References

[1] Sebastian Bader and Pascal Hitzler. Dimensions of neural-symbolic integration-a structured survey. *arXiv preprint cs/0511042*, 2005.

[2] Jadzia Cendrowska. Prism: An algorithm for inducing modular rules. *International Journal of Man-Machine Studies*, 27(4):349–370, 1987.

[3] AS dAvila Garcez, Gerson Zaverucha, and Luis AV de Carvalho. Logical inference and inductive learning in artificial neural networks. *Knowledge Representation in Neural Networks*, pages 33–46, 1997.

[4] Calvin B Harley and Robert P Reynolds. Analysis of e. coli pormoter sequences. *Nucleic acids research*, 15(5):2343–2361, 1987.

[5] Geoffrey G Towell and Jude W Shavlik. Extracting refined rules from knowledge-based neural networks. *Machine learning*, 13(1):71–101, 1993.

[6] Janusz Wnek and Ryszard S Michalski. Comparing symbolic and subsymbolic learning: Three studies. Technical report, 1992.

[7] William H Wolberg and Olvi L Mangasarian. Multisurface method of pattern separation for medical diagnosis applied to breast cytology. *Proceedings of the national academy of sciences*, 87(23):9193–9196, 1990.

System for automatic music genre recognition

Maroš Jakubec

Department of Multimedia and Information-Communication Technologies, Faculty of Electrical Engineering ,University of Zilina, Univerzitna 8215/1, 010 26 Zilina, Slovak Republic

maros.jakubec@fel.uniza.sk

Abstract. *The reason for automatic recognition of music genres is understandable. In recent years, the number of digitally recorded recordings has been increasing and new ones are constantly appearing. It is not always possible to maintain order in large databases. Due to such amount of information it is not possible to manually analyze and sort them. The processing of new recordings and their classification can be done automatically based on the audio content analysis. There exist a lot of musical genres and new ones are constantly appearing, so this work focuses only on the few of the most used ones. There is an increasing number of online music databases and user applications since the last decade. The development of basic issues has become a very fast and efficient automatic tool for classifying and searching for music content. The main goal of these tools is to acquire suitable and compact music information for representation and organization of music data. The natural way how to organize music collections is to create groups of songs in which musical aspects share similarities. Although the music genre classification is automatic, sometimes it is difficult to define the music content/genre accurately.*

The aim of this work is to explore possibilities of music genres classification from audio signal and to create a system for automatic recognition of music genres in the MATLAB programming environment. The problem solution is to summarize the applicable features as well as machine learning methods. In the theoretical part, the issue of recognition of musical genres from the audio signal is presented. The machine learning methods and acoustic features suitable for the recognition of musical genres are discussed. Based on the knowledge gained from the theoretical part, a functional system for the recognition of musical genres was created on the GTZAN dataset. We examined several classification methods including GMM, SVM, and k-NN.

Automatic genre recognition is a relatively new issue. It became known in the 1990s using various techniques and is based mainly on automatic speech recognition. Since then, a number of approaches have been tested for its solution, ranging from spectral or rhythmic analysis to searching for repeating patterns. Recognition of the genre often appears to be a trivial task for people, but that does not mean that the results are always correct or even objective. The perception of music is very subjective. Everyone has a distinctive feeling for harmony, rhythm and overall impression of music. The most widespread research in the field of automatic genre recognition has been created by George Tzanetakis, who has been working on the classification of genres for a long time. In his work he presents basic information about the given classification issue and at the same time outlines the current way of dealing with it.

Automatic recognition of music genres is one of the tasks of classifying digital signal into predefined classes. Effective and accurate automatic recognition of music genres remains a key issue. Therefore, the analog signal needs to be converted to digital form first. The music feature extraction is followed after signal processing. An appropriate combination of features must be found, so that we can describe the individual music genres as good as possible. An input to the classification function is the extracted data. The output is information about label (music genre) which is assigned to the input signal. The current challenging task of automatic search for music information is the problem of organizing, describing and categorizing music content on the Internet.

Classification is the process whereby the system determines the suitability of the tested sample for one of the defined classes. The input of the classification function is the data that are represented by the group of acoustical features. The output is then the class to which the tested sample belongs. Classification does not only mean the process of sorting input data. It is a sophisticated system of several different processes that follow each other. In our work, we implemented and tested several classification methods that include k-Nearest Neighbour (k-NN), Gaussian Mixture Model (GMM) and Support Vector Machine (SVM).

k-Nearest Neighbours (k-NN) – is one of the simplest and the oldest method of machine learning. It is a distance-based algorithm and for each vector of testing data it searches for the k nearest vectors (neighbours) from the training data. Its principle is to place the object in the space where the data is located. The object is then classified based on the affinity of most of its neighbours, so it is assigned to a class that includes most of its neighbouring objects.

Gaussian Mixture Model (GMM) – is a stochastic model that is useful in creating recognition systems. It is assumed that for each class there is a probability density function, which is expressed in the form of the mixture of components. The mixture of components is a group of individual density curves of normal probability distribution (Gaussian curves). The principle of the GMM model is therefore based on the linear combination of partial Gaussian functions. After successful model development,

the GMM method becomes an efficient, and especially fast, classification tool.

Support Vector Machines (SVM) – is another popular machine learning algorithm. This method attempts to take advantage of efficient algorithms to find a linear boundary while being able to represent highly complex nonlinear functions. The essence of the SVM method is to find such an optimal separation hyperplane on the basis of training data, which correctly divides samples into classes from the test set, which we assume to have the same probability distribution as the training data. SVM is an effective method that has been successfully used in practice, for example, for speech recognition, genre classification, and other tasks that feature a large number of attributes.

For the classification of music genres, the GTZAN database is used. Music database GTZAN serves to train models and evaluate their performance. This dataset is publicly available on the official MARSYAS software framework, created by Tzanetakis and Cook. This is a collection of musical compositions divided into 10 musical genres. Specifically: Blues, Classical music, Country, Disco, Hip Hop, Jazz, Metal, Pop, Reggae and Rock. For each genre, there is compiled a collection of 100 songs by a wide range of artists. Each recording has a length of 30 seconds, a sampling rate of 22050 Hz, and it is stored as 16 bits .au format. The whole database was divided into 2 parts in a 7:3 ratio, with a larger portion being used for training and the smaller one for testing.

The Marsyas extractor, was used to process the acoustical features extraction. Before training the model, it was necessary to standardize the input features to a range of <-1,1>. The basic selection of these descriptors was created on the basis of recommendations from authors of previous works surveying the issue of classifying musical genres, the effectiveness and difficulty of implementation. For the developed system demands, the statistical characteristics (average mean value and standard deviation) were calculated. In the proposed system, timbral textural and pitch content features were used for genre classification. Timbral texture includes the Mel-Frequency Cepstral Coefficients (MFCC) along with other spectral characteristics. For the pitch content the features extracted from Chroma are selected.

An important part of the system development is the evaluation of its classification performance. Based on this information, it is then possible to decide whether the classifier could be used or should be replaced and further modified. The evaluation of successful performance is based on how well the system is assigning the labels (name of individual genre) to new songs. F-measure generally belongs to the most used metrics for determination of the classification accuracy therefore, it was used in this work. F-measure is defined as the harmonic mean of precision P and recall R.

The program has gone through various ways of setting up classification methods. In the first part of the development, the first half of the sign data calculated for the mean value was tested. In the second part, testing was performed on all sign data, a combination of mean value and standard deviation. The purpose of this test was to find the ideal setting of the parameters of each algorithm.

In the k-NN method, the most successful classification result was achieved with the number of neighbours set to 7 and the Euclidean metric, with a success rate of 48.3 %. The best success (58 %) in the GMM method was achieved by the model setting the number of Gaussians to 32 and choosing a diagonal covariance matrix. With this classification method, the best success rate was achieved for the mean value of the signs among these three compared methods. The system achieved the best results (success rate of 51.3 %) using nu-SVC and polynomial function.

In the second part of the test, a combination of features calculated for the mean value and the standard deviation was used. After using these features, the entire classification of musical genres has improved. In this case, the k-NN method was the most successful model with the Euclidean metric setting and the number of neighbours 7. In the GMM method, the best results were obtained with the number of Gaussian 128 and the diagonal covariance matrix. For SVM, the system achieved the best results using non-SVC and linear function with a 61.2 % success rate.

The results we have achieved are not far from the most advanced systems of today. We cannot directly compare our system to other systems due to differences in experimental settings and parameters. However, these systems help us to better understand the issues so that we can improve our results in the future.

The output of the work is an application for automatic recognition of music genres, which will be accessible to users. The graphic interface for music genre recognition was created in MATLAB 2018b using App Designer. Two user interfaces have been created. The first one was designed to preview real-time music genre classification. The second was an advanced GUI version in which the user has the ability to select different parameters.

The aim of the experiment was to recognize musical genres using the k-NN, GMM, and SVM machine learning methods. The chosen classification methods achieved the best resultant success rate for the combination of features calculated for mean value and standard deviation at k-NN 69.7 %, 64.3 % for GMM and 61.2 % for SVM. Learning about the recognition of music genres has brought interesting findings and answers. To some extent, it shares some similarities with speech recognition, even though they are quite different. There is a big difference between human and machine. What may seem like a simple task to humans can be very challenging for the machine.

A huge amount of music data is nowadays on the Internet, which often makes it difficult to include songs in individual genres. Therefore, intelligent classification systems are proposed. In this paper we presented the introductory results and the concept of the system, which was created as a diploma thesis.

A simple acoustic surveillance system for child's emotion detection

Eva Lieskovská
Dept. of Multimedia and Information-Communication Technology, University of Zilina, Univerzitna 8215/1, 010 26 Zilina, Slovak Republic

eva.lieskovska@fel.uniza.sk

Abstract. *Nowadays the surveillance systems are an important part of our everyday lives. Monitoring and detection of suspicious activities or behaviors may help prevent many dangerous situations. The most of the systems are based on visual information only. However, by using of audio information we can obtain more complex view about observed situation. This could be tremendous helpful for detection of sounds produced by gun shots or abnormal audio events at public places (subway, school, etc.). The acoustic surveillance for child's emotion detection is no less important task. This paper proposes a simple surveillance system for child's cry and laugh detection. These sounds are associated with child's emotions and they are also significant indicators of child's behavior. Monitoring behaviour and safety of small children is particularly crucial in households. Having the small child under supervision can significantly help parents to take the action when it is needed. Such a cry detector can alert hearing-impaired parents, so that they can know immediately what is happening with their child. There is also a lot of publicity about children being left by parents alone in cars. These accidents may be caused by a quick errand that turns out to be longer or simple change in parents's daily routines. Therefore, it is necessary to take measures to prevent such a scenario.*

There exist numerous studies regarding the baby cry recognition. Such as surveillance application intended for child cry and laugh detection, which can make an alert to child's parents through a phone call. Or and baby cry detector in vehicles. Another example may be a system that can discriminate between different types of cries (colic, eructation, discomfort, hunger, pain, tiredness). This could be helpful for new parents or babysitters.

All of these works used different types of classification methods and audio characteristics. The brief resume of some recent works and their results is provided in our paper.

Our proposed application is based on Gaussian Mixture Model - Universal Background Model (GMM-UBM). The Universal Background Model is usually used in speaker verification systems. In the process of speaker verification, it represents person-independent characteristics which are compared to person-specific characteristics. Generally, it has shown very good performance in various audio processing tasks.

Training and testing data for our project were collected from available resources on several websites. The Audacity open source audio software was used for initial processing of audio samples. Each audio recording was converted to wav. format and downsampled to the sampling frequency 16 kHz. The annotated data for three classes was needed: cry, laugh and other sounds. The training data for the category of cry was 19 minutes long altogether. It contained 45 sounds of baby cry separately. The category of laugh contained 53 samples of baby laugh separately and it was 17 minutes long altogether. The category of other sounds (or UBM) consisted of 56 dissimilar audio samples; some of them were sounds from typical child's environment. This category was about 20 minutes long. The testing data for each category contained 20 sounds.

The audio detection task can be divided into two main parts: signal analysis, and classification.

Signal analysis- the first step of recognition algorithm consists of signal analysis in the sense of extracting some typical signal characteristics. These are usually obtained from 10-50ms frames, which may or may not overlap each other. Window function is applied on frames before calculation of features. It is used for smoothing discontinuities at the beginning and end of the sampled signal.

The system is designed in Matlab. The audio signal is divided into frames with frame length of 256 samples and 50% overlap. On each frame, Hamming window is applied in the time domain. The following 36 audio features are computed: 12 Mel Frequency Cepstral Coefficients (MFCC) without the 0th coefficient, 12 delta and 12 delta-delta coefficients.

MFCCs are obtained by applying Discrete Cosine Transform on perceptually based mel-spaced spectral band log-energies of an audio signal. To better capture spectro/temporal dynamics of the signal, MFCC vectors are usually append by their 1^{st} and 2^{nd} time derivates computed over consequent frames, known as delta and delta-delta features.

In order to increase robustness of such audio representation against noise and acoustic channel variations, MFCCs use to be also mean and variance normalized. The Cepstral mean and variance normalization (CMVN) was used in our case. By this way the data with zero mean and unit variance were obtained.

Classification - is a process that automatically assigns one of several categories or classes to an individual object represented by the audio features. In case of sound recognition the object can be an audio signal (such as various speech sounds or music tracks) and classes may be the individual speakers, types of music tracks, etc. In another words the selected classifier creates a model, from a training data set, that correctly predicts a future data (testing data set). One of the traditional and still the most popular approach of classification is Gaussian Mixture Model (GMM). It's based on probabilistic modelling of audio features distribution obtained by front-end signal analysis.

We applied two classification approaches for the recognition. In the first approach, unknown data are tested against three GMMs: cry, laugh, and UBM (which represents unknown sound or sound other than cry or laugh). The input sound data are then categorized into one of these 3 classes according to highest reached probability score. The above described approach is identification of sounds.

In the second approach - verification of sounds, GMMs for each class are built from UBM. UBM, which has been initially trained on large dataset containing various sounds, is adapted to the class chosen for recognition (e.g. baby cry) using training data of that particular class. During recognition unknown sound is evaluated by ratio of two probabilities: 1) the likelihood that a given segment of testing sound was cry or laugh respectively. 2) the likelihood that testing data belongs to the universal model i.e. the likelihood that it could be any other sound.

If the result of ratio is above threshold, audio sample is assigned to the class related to the model (cry/laugh), otherwise it was disapproved. The sound, for which the both ratios are lower than threshold (i.e. UBM fits better for the input audio data) is classified as unknown or "other sound". On the other hand if the both classifiers indicate positive classification, sound sample is assigned to the class which has a higher probability.

The optimization of threshold is an important step in sound verification algorithms. The aim is to avoid situations where system incorrectly reject target sounds or accept other sounds for the cry (or laugh). The false acceptance rate, or FAR, is defined as the ratio of the number of false acceptances divided by the number of all events. The false recognition rate, or FRR, is defined as the ratio of the number of false recognitions divided by the number of all tested events. We can balance these two characteristics to avoid false alarms or rejections by adjusting them to the point where they are equal. Their common value is called Equal Error Rate (EER). The threshold can be found in this point. The smaller value of the EER indicates the better accuracy of verification system.

At the first stage of building of the system we could see that the system detected baby cry and laugh with quite high accuracy, but the proper detection of other sounds was not satisfactory. By setting the threshold, the accuracy of another sound recognition was slightly better; however, precision of cry recognition has significantly decreased.

At the second stage of the system development the feature vectors were normalized by CMVN. At both stages we tried out a several different GMM model settings. The most accurate (71.6%) was the model with 32 Gaussian mixture components and diagonal covariance matrix. The correct predictions are balanced here and so the system will not cause too many false alarms and at the same time achieve acceptable accuracy in the recognition of cry and laugh. EER for cry–classifier was 21% and for laugh-classifier 24% while the decision threshold was set approximately to T = 1.7 for both classes. We used this model in developed application.

An application of the cry/laugh detector was developed in Matlab with AppDesigner. The application should avoid to incorrectly evaluating other sounds as cry or laugh, because it would cause an unnecessary alarm. The system measures signal power so that it does not classify silent intervals, which are below some threshold. The current version of the interface is in Slovak language.

In a final summarization we should pointed out that this project was an outcome from a diploma thesis and the aim was to present design of the system for detection of child's cry and laugh. The developed system is based on Gaussian Mixture Model - Universal Background Model approach. The classification accuracy of 71.6% was achieved. The the system performance could be improved by increasing training data and/or by adding other types of audio features.

On the importance of dropout for Checkerboard D–GEX architectures

Vladimír Kunc[1]

[1]Dept. of Computer Science, Czech Technical University, Karlovo náměstí 13, 121 35 Praha 2, Czech Republic

kuncvlad@fel.cvut.cz

Abstract. *The L1000 platform for measuring gene expression have lowered the cost by measuring only* $\sim 1,000$ *genes and then using computational approaches for reconstruction of the full profile. One of such computational approaches is the neural network D–GEX and its extension with transformative adaptive functions and checkerboard architecture. We have evaluated the influence of the dropout rate on the extended D–GEX architecture and have shown that statistically significant improvements were obtained when using dropout rate of 15% instead of the 25% that was originally used — the achieved error was 0.1278.*

Keywords

neural network, checkerboard architecture, dropout

1. Introduction

Finding out the levels of expression for various genes is very important for medical doctors and researchers (e.g., [2, 3, 6, 7]). Even though there was a price drop of the microarray technology used for measuring the gene expression in the last decade, the process is still too expensive to run a large-scale experiments. One possible solution is presented by the LINCS[1] program. The LINCS program developed a novel microarray L1000 platform based on Luminex bead technology — the L1000 platform is economically efficient because it measures only a $\sim 1,000$ carefully selected *landmark genes* and then uses computational approaches to reconstruct the expression profile of $\sim 10,000$ *target genes* [8].

There are several methods proposed for the reconstruction of the full gene expression profile; one such method is the D-GEX [1] which consists of two three layered artificial neural networks that are able to model even the nonlinear relationships. The D-GEX was further improved by the *transformative adaptive activation functions* (TAAFs) that resulted in lower prediction error [5] and also by the introduction of checkerboard architecture [4] that increases the capacity of the network and further lower the prediction error.

[1]http://www.lincsproject.org/

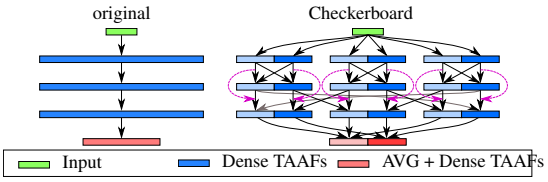


Fig. 1. The original D–GEX architectures and the checkerboard D-GEX architecture. The outputs of the towers are averaged before the output layer. Black arrows represent dense connections between blocks, purple dashed arrows represents skip connections.

The original D-GEX [1] also discussed the influence of the dropout parameter on the prediction error of the networ, however, the extended networks with TAAFs [5] and the checkerboard architecture [4] evaluated the network only with a single dropout rate due to the computational complexity of training the networks. This paper evaluates the influence of the dropout rate on the best performing network presented in [4, 5] — more specifically, we evaluate the influence of the dropout rate on the D–GEX with checkerboard architecture with 5 interconnected towers each with 4,615 neurons in each layer.

2. Methods

The evaluated network is a D-GEX with checkerboard architecture and transformative adaptive activation functions. We have used the same experimental setup as in [4].

2.1. Transformative adaptive activation functions

The TAAF is a generalization of several other adaptive activation functions that introduces four additional parameters $\alpha, \beta, \gamma, \delta \in \mathcal{R}$ per neuron that transform the original activation function $f(y)$. The output of a neuron with TAAF with inputs x_i is:

$$\alpha \cdot f\left(\beta \cdot \sum_{i=0}^{n} w_i x_i + \gamma\right) + \delta, \tag{1}$$

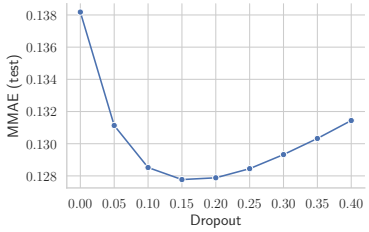


Fig. 2. The development of MMAE for different dropout rates.

where x_i are individual inputs and w_i are its weights [5]. The parameters $\alpha, \beta, \gamma, \delta$ are trained in the same manner as other weights.

2.2. Checkerboard architecture

The original D–GEX used just three dense layers, however, the number of weights between two dense layers grows quadratically with the number of neurons. Thus, there is an extension that introduced interconnected towers consisting of dense layers which allows for more neurons with the same number of weights [4]. The schema of used architecture is in Fig. 1, more details are available in [4].

2.3. Performance evaluation

Performance evaluation follows the same standards as in [4, 5]. The main metric for evaluation of the model performance is denoted MMAE and it is the absolute error averaged over individual samples and genes. We have pairwise Wilcoxon signed–rank test on the MAEs for individual samples for pairwise comparison of networks with different dropout rates.

3. Results

We have evaluated 9 different dropout rates from the interval $[0, 0.4]$ with a step of 0.05. The relationship between the MMAE on the test data and the dropout rate is shown in Fig. 2. We can observe a U-shaped curve, the networks were overfitting without a dropout; however, too high dropout rates are also harmful as those increase redundancy. The pairwise comparison using Wilcoxon signed–rank test on the MAEs for individual samples for networks with different dropout rates with significance level $\alpha = 10^{-5}$ is shown in Fig. 3. The lowest MMAE was 0.1278 when the dropout rate was set to 0.15.

4. Conclusion

Finding the optimal dropout rate might significantly influence the network performance. The original network from [4] had used only dropout rate of 25% — we have shown that the performance might be improved by using a dropout rate of 15% and that this improvement in MMAE is statistically significant at the significance level $\alpha = 10^{-5}$.

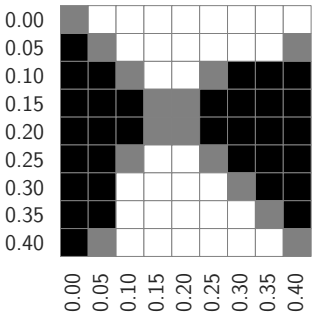


Fig. 3. Results of pairwise Wilcoxon signed–rank test on the MAEs for individual samples for different dropout rates. A cell in row r and column c is black if the model with dropout r is statistically significantly better than model with dropout c, white if worse, and grey if no statistically significant difference was observed.

Acknowledgements

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. Additional computational resources were supplied by the Ministry of Education, Youth and Sports of the Czech Republic under the Projects CESNET (Project No. LM2015042) and CERIT-Scientific Cloud (Project No. LM2015085) provided within the program Projects of Large Research, Development and Innovations Infrastructures. This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS17/189/OHK3/3T/13.

References

[1] Y. Chen et al. “Gene expression inference with deep learning”. In: *Bioinformatics* 32.12 (Feb. 2016), pp. 1832–1839.

[2] M. Hao, B. Barlogie, G. Tricot, L. Liu, L. Qiu, J. D. Shaughnessy, and F. Zhan. “Gene Expression Profiling Reveals Aberrant T-cell Marker Expression on Tumor Cells of Waldenström’s Macroglobulinemia”. In: *Clinical Cancer Research* 25.1 (Oct. 2018), pp. 201–209.

[3] A. Y. Higashi, B. J. Aronow, and G. R. Dressler. “Expression Profiling of Fibroblasts in Chronic and Acute Disease Models Reveals Novel Pathways in Kidney Fibrosis”. In: *Journal of the American Society of Nephrology* 30.1 (Dec. 2018), pp. 80–94.

[4] V. Kunc and J. Kléma. *On Tower and Checkerboard Neural Network Architectures for Gene Expression Inference*. Accepted to: *ISBRA 19*. 2019.

[5] V. Kunc and J. Kléma. *On Transformative Adaptive Activation Functions in Neural Networks for Gene Expression Inference*. Submitted to: *bioRxiv*. 2019.

[6] W.-C. Lee et al. “Multiregion gene expression profiling reveals heterogeneity in molecular subtypes and immunotherapy response signatures in lung cancer”. In: *Modern Pathology* 31.6 (Feb. 2018), pp. 947–955.

[7] J. B. Nielsen et al. “Biobank-driven genomic discovery yields new insight into atrial fibrillation biology”. In: *Nature Genetics* 50.9 (July 2018), pp. 1234–1239.

[8] A. Subramanian et al. “A Next Generation Connectivity Map: L1000 Platform and the First 1, 000, 000 Profiles”. In: *Cell* 171.6 (Nov. 2017), 1437–1452.e17.

Design of semantic model for data mining processes

Juliana IVANČÁKOVÁ[2], *Michal IVANČÁK*[1]

[1] Dept. of Cybernetics and Artificial Intelligence, FEI, Technical university of Košice, Letná 9, 042 00 Košice, Slovak Republic

[2] Dept. of Electric Power Engineering, FEI, Technical university of Košice, Letná 9, 042 00 Košice, Slovak Republic

juliana.ivancakova@tuke.sk, michal.ivancak@tuke.sk

Abstract. *It is important to think about the question: Why is data mining and knowledge discovery processes so important in today's world? Data mining is a process that allows you to collect, analyze and evaluate data. Every day is produced a large amount of data in the world. And using of data mining is growing up every day to. Sources of data we can find everywhere (healthcare, industry 4.0, IoT, e-commerce, astronomy, geography, ...). Using the CRISP-DM methodology it is possible to understand the issues, understanding and preparing data, using various methods for descriptive and predictive data analysis, evaluate and interpret data analysis results.*

Keywords

Ontology, semantic model, data mining, OntoDM, domain expert, data scientist

1. Introduction

In introduction, we look at the current problems of knowledge discovery processes. It cannot be said that we have all of them identified, but they are most relevant to us. The first problem is focused on the needs of knowledge discover processes. All processes require constant, intensive communication between a data analyst and domain expert. Data analyst needs to understand the domain of the area to be addressed, the existing relationships in it as well as the data that needs to be analyzed. On other side a domain expert must understand the results of the analysis and correctly interpret them. Nowadays, this communication is not sufficiently described, which also affects the efficiency of processes and sometimes causes unnecessary lengthy communication between them. The second problem is the fact that many tasks that a data analyst solves require repetitive writing of scripts that have the same basis, so they should be automated, saving not only time in these processes but also money in business.

2. Data mining

Data mining today is one of the fastest growing segments of business intelligence. From the statistical point of view, it is about investigating relationships or patterns in data. The point is to analyze data dependencies, identify trends, and, if that data type allows, to predict future developments. Data mining is not only a privilege for large companies, but as practice shows, these processes have a relatively large potential in smaller firms, research institutes, medicine, and so on. The process of data mining is iterative and interactive; extracted knowledge must be valid, unknown yet and potentially useful. This process is also multidisciplinary, which means that it interacts and uses methods and approaches from other scientific disciplines such as statistics, database systems, artificial intelligence, machine learning, data visualization, etc.

Fig. 1 Methodology CRISP-DM

CRISP-DM methodology (CRoss Industry Standard Process for Data Mining) is currently the most widely used approach to implement the data mining process. The goal is to provide an overview of the life cycle of the analytical project, thereby simplifying its understanding. It is based on practical experiences and consists of the six main phases

– Goal understanding, data understanding, data preparation, modeling, evaluation, deployment.

3. Knowledge technology

If we want to represent information through natural language, it may happen that this description will not be unambiguous and will not be machine readable, so it will not be able to structure them. Such a problem can be solved with the help of semantic technologies that can machine-interpret knowledge representations. These technologies include, for example, semantic networks and ontologies.

Most semantic technologies raised in the context of the Semantic Web which serves for exchange of information on the web mainly in the form of HTML pages. The semantic web architecture is represented by a layered structure which shows how the layers are dependent on each other. In the picture of semantic architecture can see at the bottom there are technologies which enable writing / defining syntax (UNICODE) that encodes the text; a URI that uniquely identifies an object. XML language provides a clear structure. RDF and ontologies represent semantic knowledge so that they can be directly machine-processed, and higher services are created over this representation than automatic proofing and deriving and securing the confidence of information by electronic signature.

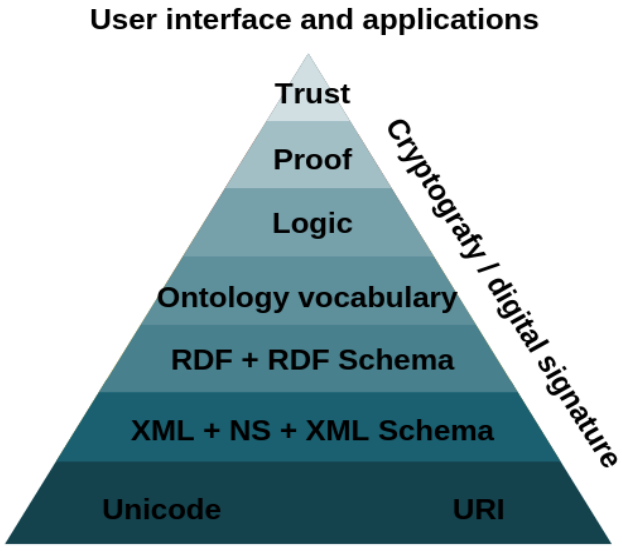


Fig. 2 The semantic web architecture

3.1 Semantic networks

One of the knowledge representations is the already mentioned semantic networks, which are characterized by the fact that within the graphical representation the objects are represented as graph nodes and the edges in the graph represent relations between these objects. RDF is used as a formalism to represent such graphs. The terms in RDF are referred to as so called unequivocally labeled URI. RDF also allows you to define anonymous resources for which a URI is generated - only refer to one RDF file, has a unique URI - unknown node. Data values can take different types of values (for example numbers, boolean values, dates, string type, ...). These types are defined in XML. The whole graph can be written as a set of triplets - Triplet = subject - predicate - object, in the direction of the arrow. The object may be a data value or a source, but the entity must be a source. So, nodes represent resources and data values, and edges represent properties. Above RDF, there is SPARQL - a query language that allows you to search for facts about the entities represented in RDF.

3.2 Ontology

The idea of using ontology in computer science is to create engineering models of reality that can be used by software, directly interpreted and justified by inferential engines. Knowledge model using classes and instances. One entity can be an instance of multiple classes. They can be organized into hierarchies from more general to more specific classes. Classes prescribe for their instances what data properties and what relationships they may have with other entities. In addition to specifying conceptualization (arranging concepts into relationships and formal notation for machine interpretability.), ontologies allow for the automatic derivation of entity facts.

3.3 RDF Schema, OWL

The basic languages for describing ontologies include RDFs and OWL. RDFS defines basic elements such as classes, properties, and instances. OWL extends the definition of classes and properties by additional boundaries, OWL is divided into 3 sub-languages. Each of these languages add more possibilities for representing knowledge (modeling), but on the other hand, the method of draining is more complex. Computationally complicated drainage, FULL provides many options, but the complexity of drainage is more difficult.

3.4 Type of ontologies

In our semantic model we will come out of the existing ontologies. OBI is ontology for describing medical processes. EXPO attempts to define the main entities needed to represent scientific experiments. It includes general knowledge of scientific experimental design, methodology and representation of results. LABORS Ontology is an extension of EXPO. Current ontologies for describing the data mining process such as OntoDM, focusing on describing methods and setting them, do not devote themselves to formalizing the data mining goal itself, and do not define a link to domain concepts describing known knowledge of the problem and data being addressed.

3.4.1 The main concepts of OntoDM

The ontology of OntoDM is based on the draft general framework for data mining. A set of basic data mining elements has been identified from this proposal [7], including:

- *dataset* - consists of data items;
- *datatype* - may be primitive or structured;
- *data mining task* - includes predictive modeling, pattern, surveying, clustering and probability distribution;
- *generalization* - output of a data mining algorithm (predictive model, pattern, clustering, probability distribution);
- *data mining algorithm* - solves the data mining task and generates summaries from data set and includes components of algorithms;
- *function* - aggregation function, prototype function, evaluation function, cost function, etc.;
- *constraint* - include assessment and language limitations;
- *data mining scenarios* - related to queries and inductive issues.

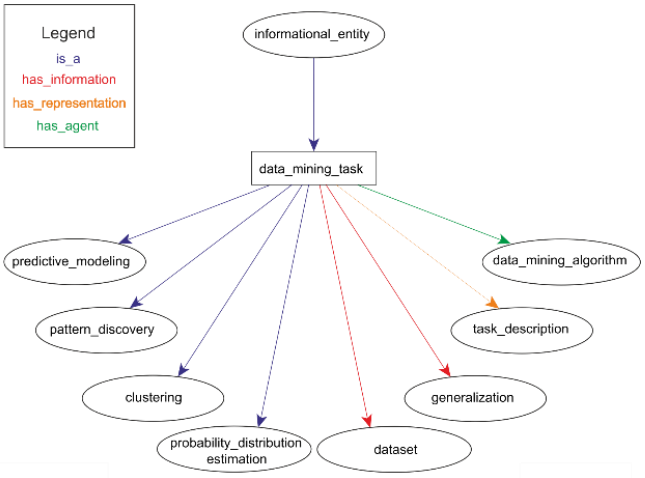


Fig. 3 Partially mapped OntoDM

4. Proposed solution

It is very important that the communication between the domain expert and the data analyst, respectively, is clear at the stage of understanding the problem and at the data understanding stage. It is not usually completely externalized and formalized when interpreting the results in the validation phase, making effective communication, sharing and reusability of the communicated knowledge impossible.

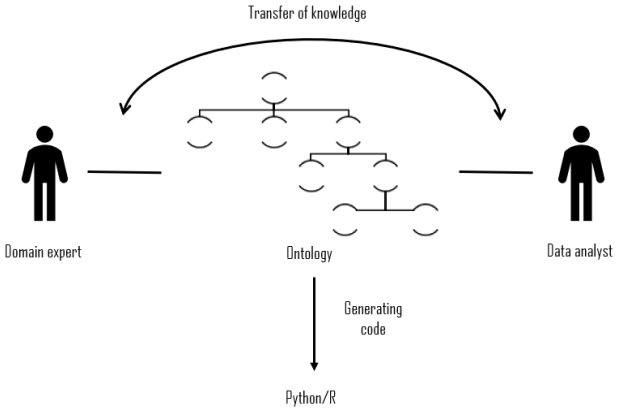


Fig. 4. Knowledge model

The proposed semantic model should consist ontologies for describe scientific research, data collected during research, various data analysis methods and a description of data analysis results that map results to specific domain concepts. The designed knowledge model is designed to simplify / improve communication between a data analyst and a domain expert, enabling them to share knowledge and reuse knowledge while ensuring data analytical reproducibility. It will also be used to support data - analytics processes automation, so that it will be able to generate preprocessing, modeling and data analysis scripts.

We will start mainly on OntoDM, ontology, which provides fundamental concepts for describing sets and data concepts and extends it to relations that allow automation to derive from the assigned task of data mining processes and description of input data, which algorithms are suitable for solving the task operation data. Also, expand ontology with metamodel to describe domain concepts.

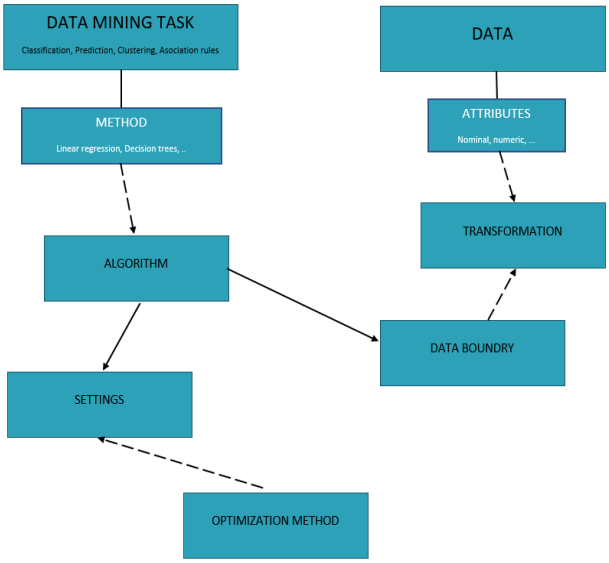


Fig. 5 Proposed solution - ontology

Acknowledgements

The research presented in this paper was supervised by doc. Ing. Peter Butka, Phd. and Ing. Peter Bednár, Phd. and was financially supported by the Slovak APVV grant SK-AT-2017-0021 and by the Slovak VEGA grant 1/0493/16.

References

[1] A. TRNKA: Využitie dolovania dát v zdravotníctve. In: *Slovak Journal of Health Sciences.* Roč. 2010, č. 1-2, s. 19-24. ISSN 1338-161X

[2] D. LAROSE: Discovering Knowledge in Data: An Introduction to Data Mining, John Wiley, 2005, 222 s. ISBN 0-471-66657-2

[3] Data Mining: What it is and why it matters, online, [cit. 2019-03-22]. Online: <https://www.sas.com/en_us/insights/analytics/data-mining.html#dmhistory>

[4] P. PANOV, L.N. SOLDATOVA, S. DŽEROSKI, Towards an Ontology of Data Mining Investigations. In: *Gama J.*, 2009.

[5] L.N. SOLDATOVA and R.D. KING, An ontology of scientific experiments. *Journal of The Royal Society Interface* 3, 795–803, 2006.

[6] D. SCHOBER, W. KUSNIERCZYK, S.E. LEWIS, J. LOMAX: Towards naming conventions for use in controlled vocabulary and ontology engineering. In: *Proceedings of BioOntologies SIG,* ISMB 2007, pp. 29.

[7] B. SMITH, W. CEUSTERS, B. KLAGGES, J. KÖHLER, A. KUMAR, J. LOMAX, C. MUNGALL, F. NEUHAUS, A.L. RECTOR and C. ROSSE, Relations in biomedical ontologies. *Genome Biology* 6, r46., 2005.

[8] [8] R.D. KING, et al.: The Automation of Science. Science 324(5923), 85–89, 2009.

[9] ANTONIOU, G., VAN HARMELEN F., Web Ontology Language: OWL, online, [cit. 2019-03-22]. Online: <https://www.math.vu.nl/~frankh/postscript/OntoHandbook03OWL.pdf>

About Authors...

Juliana IVANČÁKOVÁ was born in 1993. In 2017 graduated (MSc) at the Department of Cybernetics and Artificial Intelligence on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. At present she is a Ph.D. student in the Department of Cybernetics and Artificial Intelligence on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. She received a master's degree in Business Informatics on subject of Cancer diagnostics through suitable data mining methods. Her scientific research is mainly focused on research of Models and methods of data analysis for extraction of semantic representation of data sources.

Michal IVANČÁK was born in 1993. In 2016 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. At present he is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master's degree in electric power engineering on subject of the selectivity analysis for electrical balance of power plant with coordination with ABB. His scientific research is mainly focused on research of smart grid.

Web-based collaborative virtual environments to support cross-platform access

Marián HUDÁK[1], *Martin SIVÝ*[2]

[1] Dept. of Computers and Informatics, Technical University of Košice, Letná 9, 042 00 Košice, Slovak Republic

marian.hudak.2@tuke.sk, martin.sivy@tuke.sk

Abstract. *This article presents a concept and a prototype of multi-user interaction under web-based Collaborative Virtual Environment (CVE) to improve human training and enhance users' accesses. The introduction describes current states of virtual collaboration and its potential application. The second section focuses on CVE sharing and distribution over different network architectures. The section three specifies the proposed solution with final implementation. The conclusion summarizes the future use of CVEs.*

Keywords

Collaborative virtual environment, virtual reality, user interaction.

Introduction

Collaborative virtual environments (CVEs) mediate interaction in virtual space among more participants that may be spread over large distances. Globally, multi-user groups can participate together in one completely immersive virtual environment to achieve a desired goal. Distributed virtual environments can be purposely used as training tools for real-time 3D simulations, rescue missions, engineering and also medical courses [1].

Virtual collaboration possesses several strengths that improve its usability and demand for Virtual Reality (VR). CVE can also depict training situations that may be physically more expensive and difficult to realize than a virtual simulation [2]. Considering standard VR applications, CVEs are basically purposed to provide multi-user interaction and remote cooperation in the entire virtual space. As a result of increase in computing performance, CVEs expand with highly sensitive peripheral inputs as Inertial Measurement Units (IMUs) and haptic interfaces to achieve detailed manipulation of virtual objects [3]. These features positively affect the interaction between participants and virtual environments.

Current VR technologies are facing hardware diversity which increases demand to develop VR software which is able to operate on various systems. In order to optimize consistent CVE for all these systems, it is important to support cross-platform development. A web cross-platform can positively affect the development and progress of CVE interfaces and also simplifies the user's access through different devices [4]. Despite technological advances, there is a number of limitations related to providing multi-user interaction which have to be considered. This article approaches towards the creation of web-based CVEs focused on real-time collaboration within a large group of participants. By comparing different alternatives of sharing the virtual environment, the article proposes a CVE system for multiple users using a cross-platform application.

Shared CVE over network

Shared CVEs perform workspace collaboration irrespective of the geographical distances between participants. Because these distances between users can differ, some data consistency models have to be used to keep the data complete. In previous periods of development, two data consistency models of CVEs were shown as beneficial to use.

The first classified model for web-based CVEs is *Centralized primaries*. The communication between participants is related to the client-server architecture of sharing data through the network [5]. Distribution of data between users (clients) is mediated by the server. All of the mediated data includes user interactions and 3D object coordination, while the whole virtual CVE is fully replicated and processed by each of the clients separately. Comparing to the communication from others consistence models, the client-server can be accessed from anywhere and across multiple platforms.

The second classified is *Active replication* in which users share data with each other. Replication of data utilizes the peer-to-peer (P2P) communication without dedicated server. In P2P each of users replicates real-time data to all of the participants in the group [6]. The usage of Active replication causes significantly overloaded communication.

Comparing the client-server architecture with P2P undoubtedly proves the complexity of P2P when the number of clients is rapidly increasing. The model of Centralized primaries is much more cost-effective and independent of client's performance.

Virtual locally accessible CVE

Locally accessible CVEs are potentially useful for face to face collaboration, where each of the users is physically placed in an interactive environment with other co-users. The typical VR platform of the local CVE is the Cave Automatic Virtual Environment (CAVE) which immerses a group of participants in one physical place. Concerning CAVEs, user interaction and access are more natural than standard VR platforms.

The LIRKIS CAVE [7] at Technical University in Košice serves virtual training and educational activities through various CVEs (Fig.1). All of the users are surrounded by the same virtual environment and interact via multiple inputs [8] such as peripheral devices, smartphones, Inertial Measurements Units (IMU) and various IOT sensors. Each of the inputs can be directly adjusted to control the specific virtual object or activity in local CVE. During the virtual collaboration, all of the participants can physically interact with each other as in the real world.

Fig. 1. Locally accessible virtual collaboration in the LIRKIS CAVE.

On the other hand, there are still limitations in using locally accessible CVEs. The LIRKIS CAVE is strictly limited to place up to five participants into CVE due to its physical capacity. The collaborative activities are performed only when users are physically available. Whole remote collaboration is more complicated, due to the fact that the system operates over the LAN network.

Proposal and implementation of globally web-based CVE

In opposite to CVEs, we have decided to propose a conceptual design of the system proving CVEs on the client-server architecture. We intend to enhance CVEs for geographically remote collaboration while keeping simple accessibility and operation as global virtual collaboration. Therefore our system was named as LIRKIS Global Collaborative Virtual Environment (LIRKIS G-CVE).

The proposed LIRKIS G-CVE system (Fig.2) consists of several clients and a remote server with full-duplex data transmission. The remote server is responsible to mediate global CVE to all of the clients separately. Each client receives the entire scene and then renders only those changes that are mediated from the server. A web socket provides data transmission on both sides. The system was designed to operate over Wide Area Network (WAN) but seamlessly works through Local Network Area (LAN) with using a local server as well.

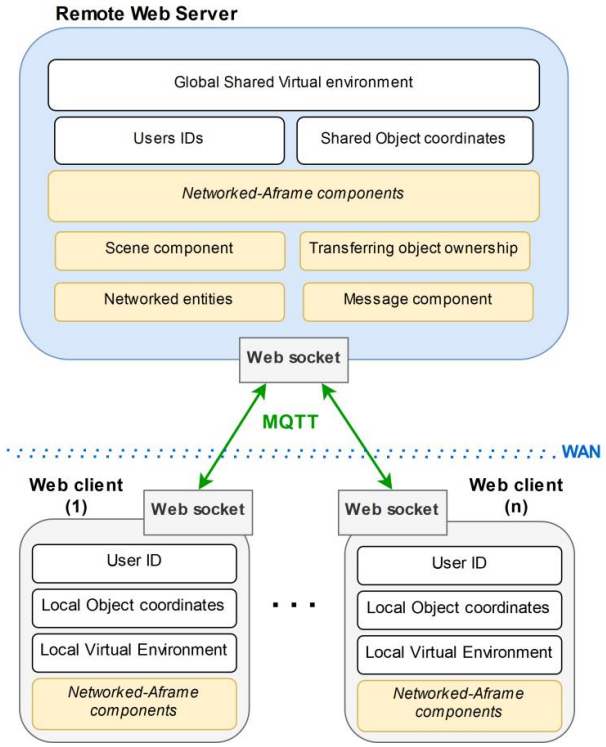


Fig. 2. The scheme of LIRKIS G-CVE web-based system.

Web-based framework A-frame

To ensure globally accessible CVE we considered utilizing VR web-based framework as necessary. To ensure cross-platform compatibility, the A-frame [9] technology (open-sourced) was chosen to build a prototype of the system. The A-frame supports fully immersive web-based VR founded on the entity-component architecture with unlimited access to other web technologies. The A-frame utilizes JavaScript and HTML to extend its support into many fields of web VR. The visualization core of the A-frame is based on the Three.js library for rendering interactive 3D and 2D graphics within any compatible web browser. To deploy final LIRKIS G-CVE the Glitch project tool was chosen with respect to run online global CVE under public URL.

Remote web server

In the stage of remote server implementation, we considered ***Node.js*** and ***Networked-Aframe (NAF)*** as stable technologies to rapid prototyping of web-based CVE. Both of them are open-sourced and aim to server-side scripting with utilizing JavaScript language. Prior to other frameworks, the Node.js offers wide scalability to various data

structures transferred between client and server [10]. Utilization of Node.js is beneficial to provide real-time assurance of parallel connections [11]. This feature positively affects the process of the asynchronous client-server communication. To handle various clients' connections was important to provide the option of non-blocking their I/O. In terms of sharing web-based CVE, several NAF components were deployed to the remote server.

The ***NAF components*** represent a set of utilities improving CVE's access and interaction. Based on virtual collaboration in the LIRKIS G-CVE, it was exploited the following ones:

- networked scene component,
- networked entities,
- transferring object ownership,
- message component.

Networked scene component

The Networked scene component contains a globally shared space through the network. All of its parts represent modules such as the server configuration, callback function, the virtual room, and network adapter. The server configuration utilizes URL and port of remote WebSocket over which the server is operating. After the client is successfully connected to the server, the callback function triggers continuous network communication. The virtual room provides shared CVE represented by a 3D static object as referential to all of the clients. By using network adapters, it is possible to serve client-server (WebSocket) or peer-to-peer (WebRTC) communication.

Networked entities

The ***Networked entities*** represent virtual avatars and objects which can be manipulated by multiple clients. By means of networked entities, all of their properties such as position, orientation, scale, and texture are synced among all of the clients in real-time. The server mediates each change of entities and distributes them directly to all clients. All of the CVEs can be enriched of various virtual objects within their dynamic behavior.

Transferring object ownership

Specific objectives of remote virtual collaboration closely relate to distributing interaction among users and virtual object. A key solution focuses on the ***Transferring object ownership*** among multiple clients during virtual collaboration. Each user can be an owner of virtual object. The owner is able to manipulate virtual object and change its coordinates while other users can only spectate until the ownership is finished. Afterwards, the other user can take the ownership.

Message component

Sending messages between users utilizes the ***Message NAF component*** containing the clients' time logs, messages, and the number of logged participants. While the previous Message NAF component provided simple multi-user messaging, our effort was focused on storing all of the data during the virtual collaboration. Therefore it was performed the extension of recording logs of the client-object interaction. This feature can be helpful to evaluate clients' activity after virtual collaboration is terminated.

Elimination of data stream

A key question in the data stream is how it is possible to provide stable remote connection and minimalize the amount of data packets. The stability of remote connection during virtual collaboration is strictly affected by network capabilities. In this case, we assumed the elimination of communication lacks by using MQTT protocols between clients and server. The MQTT minimizes network bandwidth and provides stable data transmission. In other words, the MQTT is preferred in systems, which distribute data among large groups of end-devices connected into the network.

Client interface

By means of A-frame, the client interface is fully adaptable according to the end-device that is running on (Fig.3). Each of end-devices has typical features, which can be used as primary to control CVE. The A-frame guarantees standard support of using smartphones, desktops, tablets and various types of Head Mounted Displays (HMD). In the case of peripherals, all of inputs can be obtained by utilizing JavaScript and A-frame entity-components. In the process of rendering CVE, the A-frame operates with the same NAF components included in the server. To ensure user's spawning into the CVE there was implemented AJAX client script. The script queries the remote server to create a virtual avatar and coordinates of its position where it places in CVE.

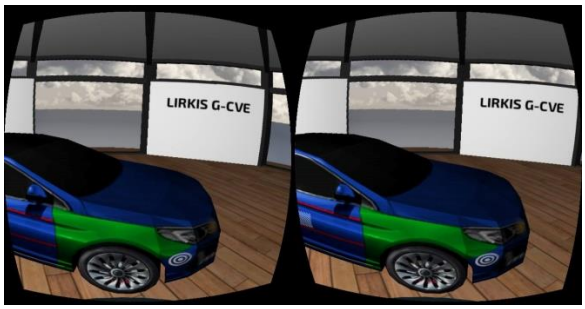


Fig. 3. Visualization of the G-CVE on smartphone.

Peripherals and interaction

Standard peripheral devices were directly supported by built-in A-frame features to operate with keyboard and mouse. Despite the internal support of these facilities, we focused on data acquisition through web browsers JavaScript API to access various peripherals. This option was beneficial for obtaining different input data from various types of connected devices.

Currently supported peripheral inputs includes:

- touch,
- mouse and keyboard,
- controllers with three degrees of freedom,
- VR gamepad and joystick.

However, handling of multiple peripheral accesses can overload the client-server data transmission. Although each client operates at most three devices simultaneously, a large group of users in CVE can cause server failure. In order to properly minimize network data transmission (Fig.4), the client sends to the server only virtual avatar coordination and state of virtual ownership called *avatar data*. The avatar data is contained in the simple structure which can be quickly processed by the server. After the server obtains *avatar data*, it resends it to all clients synchronously.

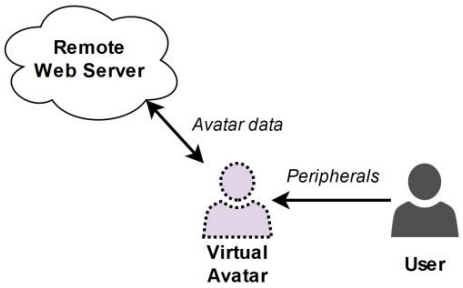


Fig. 4. The scheme of peripherals and Avatar data transmission.

Future work and training

In the current state of implementation, the system provides only simple interaction between various sized users groups. As shown in the Fig.5, trainings are focused on visuospatial navigation and space coordination control. During training, participants use smartphone headsets making them more immersed. In future work, we assume that CVE could be used in complex trainings and for educational purposes. Therefore, we consider it important to test the system with more connected users, in future research. We are planning to extend system's control in a manner that it won't be limiting for human. In this intention, we are considering to design a quickly adaptable environment.

Fig. 5. Visualization of prototyped CVE system.

Conclusion

In this paper we have presented cross-platform web-based CVE system to enhance multi-user interaction and training. Distribution and sharing of CVEs are positively impacting communication and training among a large group of users. This assumption forms a motivation to create immersive CVEs with much more convincing responses. Our goal is to rescale usability an improve user access to remote virtual environments through cross-platforms. We suppose that the system will be fully capable of providing fast development of virtual environments and their functionality. Thanks to A-frame scalability to provide cross-platform CVEs with multiuser interaction.

Acknowledgements

This work has been supported by the APVV grant no. APVV-16-0202 "Enhancing cognition and motor rehabilitation using mixed reality".

References

[1] Sharma, S., Devreaux, P., Scribner, D., Grynovicki, J., & Grazaitis, P. Megacity: A Collaborative Virtual Reality Environment for Emergency Response, Training, and Decision Making. *Electronic Imaging*, 2017, 2017.1: 70-77.

[2] Alatalo, T., Koskela, T., Pouke, M., Alavesa, P., & Ojala, T. VirtualOulu: collaborative, immersive and extensible 3D city model on the web. In: *Proceedings of the 21st International Conference on Web3D Technology*. ACM, 2016. p. 95-103.

[3] Hudák, M., Sobota, B., & Korečko, Š. (2018, November). Gesture Control for Cognitive Training Based on VR Technologies. In 2018 16th International Conference on Emerging eLearning Technologies and Applications (ICETA) (pp. 209-214). IEEE.

[4] Paiva, P. V., Machado, L. D. S., Valença, A. M. G., De Moraes, R. M., & Batista, T. V. Enhancing collaboration on a cloud-based CVE for supporting surgical education. In: *2016 XVIII Symposium on Virtual and Augmented Reality (SVR)*. IEEE, 2016. p. 29-36.

[5] Caetano, C., Knechtel, M., Resmini, R., Garcia, A. C., & Montenegro, A. Simulated architecture and programming model for Social Proxy in Second Life. In: *2010 International Conference on Information Society*. IEEE, 2010. p. 31-35.

[6] Paiva, P. V., dos Santos Machado, L., & Batista, T. V. V. A collaborative and immersive VR simulator for education and assessment of surgical teams. In: *2015 XVII Symposium on Virtual and Augmented Reality*. IEEE, 2015. p. 176-185.

[7] Hudák, M., Sobota, B., & Korečko, Š.. On architecture and performance of LIRKIS CAVE. *In 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, Debrecen, 2017, pp. 295-300.

[8] Hudák, M., Korečko, Š., & Sobota, B. On architecture and performance of LIRKIS CAVE. *In Open Computer Science* 8.1, 2018, pp. 1-9.

[9] Aframe: A web framework for building virtual reality experiences homepage (2019), *https://aframe.io/*

[10] Tilkov S, Vinoski S (2010) Node.js: using JavaScript to build high-performance network programs [J]. IEEE Internet Computing 14(6), pp.80–83

[11] Fang, D., Xu, H., Yang, X., & Bian, M. (2019). An Augmented Reality-Based Method for Remote Collaborative Real-Time Assistance: from a System Perspective. *Mobile Networks and Applications*, pp.1-14.

About Authors...

Marián HUDÁK born in 1992 in Košice, Slovakia. In 2017 he graduated (MSc) at the Department of Computers and Informatics of the Faculty of Electrical Engineering and Informatics at Technical University of Košice. He defended his master's thesis in the field of Informatics. Currently, he is a PhD student in the same department. His research is focused on virtual reality, gesture recognition, and collaborative immersive technologies.

Martin SIVÝ born in 1993 in Humenné, Slovakia. In 2017 he graduated (MSc) at the Department of Computers and Informatics of the Faculty of Electrical Engineering and Informatics at Technical University of Košice. He defended his master's thesis in the field of Informatics. Currently, he is a PhD student in the same department. His research is focused on virtual reality and smart user interfaces.

Entity Recognition Using Contextual Embeddings

Jan PICHL[1]

[1]Dept. of Cybernetics, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

pichljan@fel.cvut.cz

Abstract. *In this paper, we present a Named entity recognition sequence labeling task using contextual embeddings such as ELMO or BERT. We compare the result using traditional BiLSTM or BiLSTM-CRF models using word embeddings with the approaches taking advantage of contextual embeddings. These embeddings are trained on large corpora which helps the model to understand the language even if the task-specific dataset is limited. Additionally, the contextual nature of the representation allows us to describe the same word with a different representation regarding the context. For that purpose, we test the models on a commonly used dataset CONLL 2003 and a relatively small in-house-labeled dataset of conversations between bot and a user.*

Keywords

natural language understanding, natural language processing, entity recognition, entity disambiguation

1. Introduction

Entity recognition also called Named entity recognition (NER) is one of the key tasks in Natural language understanding (NLU). The goal of the task is to find a span of words representing the entities in a given sequence of tokens. Additionally, we usually need to determine the type of entity recognized in the text. This task can be used in various systems dealing with the natural language such as question answering or dialogue systems. There are several approaches to solve this task including sequence labeling, knowledge base matching or usage of gazetteers.

In this paper, we are focusing on the sequence labeling method which works only with the information given in the input text with no external knowledge. This approach allows finding a span of words representing entities base on the sentence structure. Additionally, the type of the entity can be recognized if the information about the type is contained in the sentence (e.g. *Let's chat about the Matrix movie*). On the other hand, if the word *movie* is omitted, the type of the *Matrix* entity cannot be determined without external knowledge.

2. Task Description

The goal of the NER task is to given a sequence of words $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ assign a sequence of tags $\mathbf{t} = (t_1, t_2, \ldots, t_n)$, one tag for each word. We use a *inside, outside, begin* (IOB) [1] schema commonly used in this task. Additionally, we need to predict a type of the entity (e.g. *B-LOC* - location). The types are domain specific.

The named entity is defined as a real-world object [2], physical or abstract, for example, organizations, locations, people. Typically, these entities have their representations in knowledge bases on which they can be mapped. We do not require for the entity to have this representation which allows us to recognize, e.g., *my life* as an entity in the sentence *Letś talk about my life.* It is useful for example in dialogue system where we need to know what a user wants to talk about.

3. Datasets

We use two datasets to test the NER models. The first one is commonly used CoNLL 2003 dataset [3] and the second one is an in-house dataset of manually annotated dialogue conversations between a human and bot Alquist [4].

The CoNLL 2003 dataset is commonly used to compare NER systems. It contains four types of entities: LOC (location), ORG (organization), PER (person), MISC (other).

Tab. 1. Number of samples in splits of CoNLL 2003 dataset.

Split	Sentences	Tokens
Training	14,987	203,621
Testing	3,684	46,435

The Alquist dataset is a small set of dialogue utterances with 12 types of entities (e.g. movie, sport, song). The data were randomly selected from the conversation logs.

We selected these two datasets to show the contribution of contextual embeddings especially if the number of samples is low.

Tab. 2. Number of samples in splits of Alquist dataset.

Split	Sentences	Tokens
Training	2,998	23,374
Testing	524	4,085

4. Experiments

We have experimented with 5 different models: (1) BiLSTM [5] with fasttext embeddings [6], (2) BiLSTM-CRF [7] with fasttext embeddings, (3) BiLSTM with ELMO embeddings [8], (4) BiLSTM-CRF with ELMO embeddings and (5) pre-trained and fine-tuned small uncased BERT [9] model.

4.1. Metrics

The input sentences typically contain a few entities and much more tokens which are not entities. Because of that, the usage of accuracy as the main metric would be biased by this imbalance. Therefore, we report F_1 score which is a harmonic mean of precision and recall.

$$p = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}$$

$$r = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}$$

$$F_1 = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

4.2. Models

AS the input representation of the words, we use the following embeddings: 300-dimensional pre-trained fastText and 1024-dimensional contextual ELMO. The BERT model is a fine-tuned uncased 768-dimensional version. For the bidirectional LSTM layers, we use 128 cells in each direction. Non-CRF version is followed by a classification dense layer with softmax activation, and the CRF version is followed by a 128-dimensional dense layer and CRF layer.

Table 1 and Table 2 show the results of the individual models list in the order as described at the beginning of this section. We list the F_1 score of each model as well as the number of parameters, resources consumption and training time. Please note that the asterisk sign in the last row means that the number of parameters contains both trainable and fixed parameters.

Tab. 3. Results of entity recognition on CoNLL 2003 dataset.

Model	F1 score	# Params	GPU memory/ power usage	Time	Memory usage
(1)	0.83	**441,866**	39/6%	**7min 1s**	**181.65 MB**
(2)	0.84	473,602	36/5%	8min 21s	192.98 MB
(3)	0.87	1,183,246	90/49%	8min 36s	1700.14 MB
(4)	**0.88**	1,183,366	88/51%	11min 20s	2273.50 MB
(5)	**0.88**	12M*	46/91%	18min 25s	3,682 MB

Tab. 4. Results of entity recognition on Alquist dataset.

Model	F1 score	# Params	Avg. GPU memory/ power usage	Time	RAM usage
(1)	0.76	**441,866**	39/6%	**2min 49s**	**153 MB**
(2)	0.76	473,602	36/5%	2min 24s	181 MB
(3)	**0.82**	1,183,246	90/49%	2min 50s	1130 MB
(4)	**0.82**	1,183,366	88/51%	3min 37s	1302 MB
(5)	**0.82**	12M*	44/84%	8min 24s	3,555 MB

5. Conclusion

We experienced with several models for the NER task. The experiment results show that the contextual embeddings improve the performance significantly especially if the training data does not contain a lot of examples. On the other hand, usage of the contextual embeddings requires more computation resources (both RAM and CPU/GPU), and the training process is more time expensive.

References

[1] Ramshaw, L. A.; Marcus, M. P. Text chunking using transformation-based learning. In *Natural language processing using very large corpora*, Springer, 1999, pp. 157–176.

[2] Nadeau, D.; Sekine, S. A survey of named entity recognition and classification. *Lingvisticae Investigationes*, volume 30, no. 1, 2007: pp. 3–26.

[3] Sang, E. F.; De Meulder, F. Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition. *arXiv preprint cs/0306050*, 2003.

[4] Pichl, J.; Marek, P.; Konrád, J.; et al. Alquist 2.0: Alexa Prize Socialbot Based on Sub-Dialogue Models.

[5] Chiu, J. P.; Nichols, E. Named entity recognition with bidirectional LSTM-CNNs. *Transactions of the Association for Computational Linguistics*, volume 4, 2016: pp. 357–370.

[6] Bojanowski, P.; Grave, E.; Joulin, A.; et al. Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, volume 5, 2017: pp. 135–146, ISSN 2307-387X.

[7] Huang, Z.; Xu, W.; Yu, K. Bidirectional LSTM-CRF models for sequence tagging. *arXiv preprint arXiv:1508.01991*, 2015.

[8] Peters, M. E.; Neumann, M.; Iyyer, M.; et al. Deep contextualized word representations. *arXiv preprint arXiv:1802.05365*, 2018.

[9] Devlin, J.; Chang, M.-W.; Lee, K.; et al. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

Eliminating Semantic Ambiguities Using Intensional Logic

Zuzana BILANOVÁ[1], Matúš UCHNÁR[2], Martin ŠTANCEL[3]

[1,2,3]Dept. of Computers and Informatics, Technical University of Košice, Letná 9, 042 00, Košice, Slovak Republic

zuzana.bilanova@tuke.sk, matus.uchnar@tuke.sk, martin.stancel@tuke.sk

Abstract. *The paper explains why it is extremely difficult to interpret the meaning of the natural language expressions. In the first part of the paper, the transparent intensional logic is described as a method that allows to reveal ambiguities in the natural language and formalize its meaning in the logical constructions. Consequently, the disadvantages of this method are pointed out, as well as the way to overcome them for future effective use in the computer science.*

Keywords

Logical analysis of natural language, predicate linear logic, transparent intensional logic.

1. Introduction

For the speaking and understanding spoken, people use not only words of language but also patterns of language sentences. Linguists call these patterns language rules, and their full list is called mental language grammar [1]. It is possible to assign meaning to the words of a language, but it is not true that by assigning meanings to individual words of the sentence, it is possible to determine its meaning (in the sentence *the lady with a dog that Mr. Smith likes*, is not clear whether Smith like a lady or her dog) [2]. To give some information to another person, it is necessary to create a mental model of the information, translate it into the mental grammar, and select and organize the form of the words that allows the listener to reconstruct the original mental model [3]. A similar process is necessary for the computer processing of the meaning of natural language expressions, while the meaning must be revealed by analyzes. Logic Analysis of natural language (LANL) has multiple methods. This article focuses on one of them - transparent intensional logic.

2. Transparent intensional logic

Pavel Tichý [4], the creator of transparent intensional logic (TIL), specified ramified theory of types where the basic type level contains the simplest natural language entities. Tichý extended Church's simple theory of types, which contained sets of individuals and truth values, by a time-spatial component. The object base B v TIL contains set of truth values o, set of individuals ι, set of time points τ and set of possible worlds ω.

The hierarchy of types in TIL allows to construct objects above the object base B as a function above that base [5]. (α-)intentions are objects of type $(\alpha_{\tau\omega})$ that representing functions from possible worlds at time points to objects of type α. TIL intentions (proposition, individual roles, object properties, relations between objects...) can be expressed by using predicates.

In TIL, the meaning of expression is expressed by using abstract algorithm-structured procedures called constructions. Constructions may be variables (x) - construct object dependent on its valuation, trivialization $({}^0X)$ - constructs object X without any change, closure $(\lambda x_1, ..., x_n X)$ - constructs the function, or composition $([XY_1, ..., Y_n])$ - constructs the result of an application of function to arguments.

Tichý's TIL, based on our analysis [6] [7], is the most expressive mechanism of LANL. Its disadvantage (the impossibility of specifying a complex deductive calculus, the impossibility of creating a fully automated implementation, etc.) can be removed by linking TIL to another logical system, usually not used in LANL. The best choice is to choose a logic system that would increase the expressive power of TIL [8]. Such a system could be predicate linear logic.

3. Predicate linear logic

Linear logic was formulated by Jean-Yves Girard in 1987 as a formalism describing problems with an explicitly determined amount of resources. Since intensions in the natural language can be described by using predicates, we need to introduce predicate linear logic (PLL) [9]. Syntax of the PLL formulas φ and terms t can be expressed by following production rule in Backus-Naur form [10]:

$$\varphi ::= 1 \mid 0 \mid \bot \mid \top \mid P(t, ..., t) \mid \varphi \otimes \varphi \mid \varphi \& \varphi \mid \varphi \oplus \varphi \mid \varphi \parr \varphi \mid \varphi \multimap \varphi \mid \varphi^{\bot} \mid !\varphi \mid ?\varphi \mid (\forall x)\varphi \mid (\exists x)\varphi \tag{1}$$

$$t ::= x \mid c \mid f(t, ..., t) \tag{2}$$

where $1, 0, \bot, \top$ are logical constants, $P(t, ..., t)$ is predicate, $\otimes, \parr$ are multiplicative logical connectives, $\oplus, \&$ are additive logical connectives, $\multimap$ is linear implication, $(.)^{\bot}$ is negation, $!, ?$ are modal operators, $\forall, \exists$ are quantifiers, x is variable, c is constant and $f(t, ..., t)$ is the application of a functional symbol to a term t.

The properties of PLL (PLL constants can capture not only the truth but also the meaning of the expressions in the form of predicate linear formulas, PLL has four logical connectives, because of which it is an extremely expressive system, PLL operators describe depended and independent choice of individual, PLL can describe the consumption of resources, what is a special way to capture changing world conditions, PLL can be automated due to a fully-defined deduction calculus) [11] represents the indisputable advantages that traditional LANL systems do not provide.

4. Conclusion

In this paper, LANL was described as a tool to detect semantic ambiguities. The goals of the presented research is possible to divide into two parts, where each represents an original contribution in the processing of the natural language. The first phase will be based on existing and already published results in which TIL has been identified as the logical system for LANL with the highest express power. Because of several serious disadvantages an original logical system of source-oriented character will be created. Its ambition will be overcome mentioned disadvantages. Following the first phase of the project, its capabilities will be tested and compared with existing solutions not only through exact logical proofing, but also through its implementation of a semantic machine. The implementation of a semantic machine, based on the natural language translation algorithm in logical constructions, represents the practical part of the presented research.

Acknowledgements

Research described in the paper was supervised by Prof. Liberios Vokorokos, FEI TUKE in Košice and supported by the following projects: Faculty of Electrical Engineering and Informatics, Technical University of Košice under the contract No. FEI-2018-59: *Semantic Machine of Source-Oriented Transparent Intensional Logic* and Slovak Research and Development Agency under the contract No. SK-AT-2017-0012: *Semantics technologies for computer science education*.

References

[1] CALVIN, W. H. *How Brains Think: Evolving Intelligence, Then And Now*. 1st ed. Basic Books, 1997.

[2] JACKENDOFF, R. *Patterns In The Mind: Language And Human Nature*. 2nd ed. Basic Books. 1995.

[3] BICKERTON, D. *Language and Human Behavior*. 2nd ed. University of Washington Press, 1996.

[4] TICHÝ, P. *The Foundations of Freges Logic*. 1nd ed. Berlin and New York: De Gruyter. 1988.

[5] RACLAVSKÝ, J. On partiality and Tichs transparent intensional logic. *Magyar Filozfiai Szemle*, 2010, vol. 54, p. 120 - 128.

[6] BILANOVÁ, Z., UCHNÁR, M. Comparison of the approaches of Montague and Tich within a logical analysis of an English sentence. In *POSTER 2017*. Prague (Czech Republic), 2017, p. 1 - 6.

[7] COCCHIARELLA, N. Richard montague and the logical analysis of language. *G. Floistad (ed.), Contemporary Philosophy: A New Survey I*, 1981, vol. 2, p. 113 - 154.

[8] BILANOVÁ, Z. Increasing the expressive power of the intensional logic used for logical analysis of the natural language sentences. In *SCYR 2018: 18th Scientific Conference of Young Researchers*. Koice (Slovak Republic), 2018, p. 128 - 129.

[9] DEMETEROVÁ, E. The role of predicate linear logic in computer science. *SCYR 2014: 14th Scientific Conference of Young Researchers*, 2014, p. 273 - 274.

[10] VOKOROKOS, L., BILANOVÁ, Z., MIHÁLYI D. Linear logic operators in transparent intensional logic. In *Informatics 2017. - Danvers : IEEE*. Koice (Slovak Republic), 2017, p. 420 - 424.

[11] BILANOVÁ, Z. Source-oriented approach in logical analysis of the natural language. *Electrical Engineering and Informatics 9 : proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. - Košice : FEI TU*, 2018, p. 627 - 630.

About Authors...

Zuzana BILANOVÁ graduated from a Technical University in Košice with a major in informatics and is currently in the third year of the postgraduate study. She has been focusing on logical analysis of natural language sentences, nontraditional logic systems in computer science and the theory of types.

Matúš UCHNÁR graduated from a Technical University in Košice with a major in business informatics and is currently in the fourth year of the postgraduate study. His thesis is aimed towards malware analysis and he has also been interested in creating the semantic machine based on transparent intensional logic.

Martin ŠTANCEL graduated from a Technical University in Košice with a major in informatics and is currently in the second year of the postgraduate study. He has been focusing on machine learning and he has also been interested in creating the semantic machine of transparent intensional logic.

Uniformed Intelligent user interface for Smart Home using Virtual Reality

Martin SIVÝ[1], *Marián HUDÁK*[2]

[1,2] Dept. of Computers and Informatics, Technical University of Košice, Letná 9, 042 00 Košice, Slovak Republic

martin.sivy@tuke.sk, marian.hudak.2@tuke.sk

Abstract. *Nowadays we live in a time that is largely driven by technology. In this paper we briefly introduce our vision and progress with Microsoft HoloLens and IoT SMART home. Article is focused on Intelligent user interface which is presented to user using device Microsoft HoloLens. With the help of these technologies and appropriate learning methods it is possible to improve our day to day life, even for those who are disabled. As a whole this process can be performed in real life or prototyped virtual environment and therefore be safer in the process of learning how to use it for disabled people. Prototyping in virtual reality (VR) can also provide early advantage of confirming the correctness of placement IoT devices in SMART environment.*

Keywords

virtual reality, Microsoft HoloLens, SMART home, user interface.

1. Introduction

The field of virtual reality is widely spread in today's era, and its application is highly up to date. The elements of smart environment can be divided into two main categories and those are actuators and sensors [1]. In order for devices and sensors in an intelligent environment to be user-controllable, users needs to be informed about how to control them through the user interface. Personal laptops (smartwatches, smartphones, glasses, ...) [2] are highly available to the user and should therefore be an ideal option for easy interaction with devices in a smart environment. The current state of user interfaces is focused on creating user interfaces for mobile devices and adapting everything to support small displays. However, the intelligent environment should also support the monitoring of natural movement of a person, such as limb gestures. The environment in which people live can be controlled via the position of body in space, gestures, speech, or writing. Thus, the user interface that is inspired by natural behavior has a great chance to become easily understandable for a large group of people ranging from children to retirees.

User interface will be provided using Microsoft Hololens as a mixed reality device. With the use of such technology user can freely move through rooms. There are two ways how we can describe the interconnection of SMART environment and Virtual reality:

- SmE is superior and VR is a complementary part of it as in(Fig. 1.).
- Virtual reality is master and SmE is virtualized only.

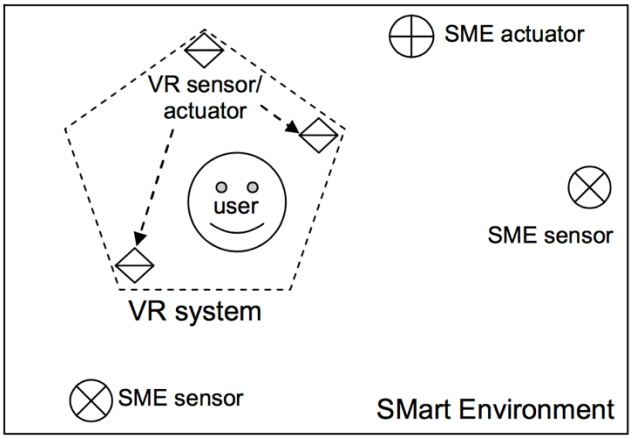


Fig. 1. *Conceptual model: SMART environment master – Virtual Reality slave [1]*

2. IoT and HoloLens

Both Hololens and IoT devices are connected to the internet. Used technology concept is based on server side architecture for sharing data between IoT and HoloLens. Therefore HoloLens and Cloud Web Server have unifying standard for data structure. Web server is running on Java Spring framework using RESTful API. New IoT devices can be configured or removed from the server at any time.

Applications developed for HoloLens can analyze surfaces and place virtual objects on them. Such virtual objects can be arbitrarily deployed into free space in the real

world and can also represent IoT devices. Applications that are not customized to 3D holographic mode work in 2D mode as a window. The window can be moved in 3D space and be placed, for example, on a wall. Thus, the user always has an interface with the current data that can pre-configured to user's needs.

3. User interface for Mixed reality

It is always necessary to design a user interface in such a way that it satisfies users' needs, provides rich virtual reality experience and is user-friendly. We want to focus on unifying all approaches in the field of user interaction with SMART home. User will have only one "application" which will integrate all devices using data from Cloud Web Server. User will work with 3D objects or 2D windows. Although it is possible to see several windows simultaneously in space, user can only interact with one at a time. Those with which the user is not currently working are switched to sleep mode to achieve limited memory usage but they will be always available when initialized. Control is intuitive with hand gestures, movement and head positioning [3], and no other input devices are required. In some cases, the lack of object visibility may also be a disadvantage. At distances above 3 meters, objects tend to change slightly and they may jump [4]. On the contrary, at a distance less than 0.85 meter, user is no longer able to see the object he is facing. User interface in the room can be seen in (Fig. 2.).

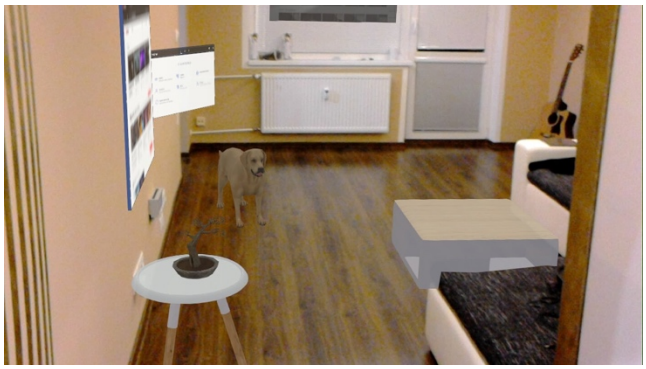


Fig. 2. *Visualization of Holograms and 2D windows with user interface through Microsoft HoloLens*

4. Three main parts of project, IoT sensors, HoloLens and server side application

This work focuses on unified application and communication interface among IoT and Hololens as well. Communication can be divided between three independent parts and those are IoT, Hololens and server side of the application. Devices communicate only with the server and it propagates data further. Therefore more than one Microsoft Hololens can be connected to the environment, since all devices are managed by the server and more than one device can be used for multiple users in the household (Fig. 3).

Server can store last information about devices and can check their state without the direct request from user. Server part of application communicates in form of JSON messages and it was implemented in JAVA language with the use of Spring boot framework. This enables the use of different Cloud services. IoT devices collect local data from sensors such as movement sensor, thermometer, thermostat, lights and others. This data and internal state of devices can be sent with their specific ID to the server. Devices can receive responses with instructions about their further behaviour. Server has a PostgreSQL database, where last states of IoT devices are stored, so that it can manage these information and propagate them to Microsoft Hololens devices.

Server stores all changes of IoT devices as time progresses for later analysis of behavior that takes place in household. Neuron network could be used to analyze and use this data for machine learning and to suggest the set-up of household to the user that would reflect his needs in the best way or that could provide the most used options for different devices in household the way he uses them in some context menu. This could make the user experience easier and user would not need to search for the action he wishes to execute for a longer time. Other options could be stored in for example a separate menu or the list of functions could be ordered according to user's frequent history of their use.

Server application can have separate privileges apart from user's Hololens to control the household without his input independently. It could decide how to control IoT devices according to recently measured data from sensors or user could asynchronously plan the behaviour of household in advance or set up boundary values for certain actions to take place.

The last part of proposed system is an application for Microsoft Hololens. This application visualises and controls different IoT devices that were registered in environment by server. User can walk through the household with the Microsoft Hololens on his head and interact with the IoT devices. Microsoft Hololens needs to be connected to the server and have Internet connection. User can invoke virtualised menu when looking at any device and choose an action provided by the environment based on the device's properties, such as turning the lights on or off, setting up desired temperature on thermostat or read the current one or schedule future actions. Unified user control of more complex devices through one API is the main goal of such system.

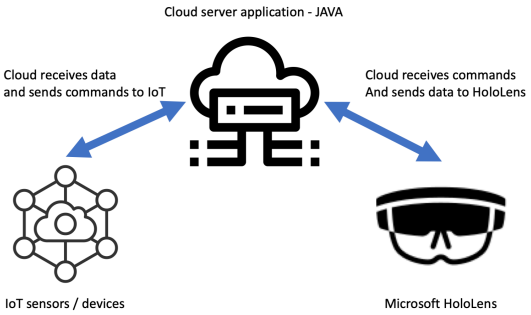


Fig. 3. Communication design among devices

5. CAVE system for prototyping SMART environment

Since the interface is made uniformly, it can also be used in the CAVE system environment. The CAVE system is particularly useful for its ability to immerse user and also user can feel more natural in such environment than using other HMD devices or VR by smartphones or other headsets. Also, there can be more users in CAVE at the same time and they can interact with one another. In this environment users can cooperate with other users and objects. One such CAVE system is also in LIRKIS [6] Technical University in Košice Fig. 4.

Fig. 4. *CAVE in LIRKIS – fully immersive interface*

It is possible to prototype IoT devices for SMART households rapidly with the use of CAVE environment. One can use possibilities such as creating 3D model of a home, placing virtual IoT devices inside of it, manipulating them freely, deleting or copying them as desired, and others. These virtual IoT devices can simulate devices in real world and send data to server side of application with their actual state. Server can communicate with them via CAVE interface. Server is completely separated from CAVE system and Hololens, only IoT devices communicate with server.

Another great advantage of using CAVE system for prototyping is endless availability of IoT devices and sensors. They can be programmed for a simulated environment and can send approximately generated data to server.

Hololens user interface can be replaced and omitted when fast prototyping is desired. CAVE systems and its tracking device Optitrack [8] can be used instead of this headset. User can have a marker on himself that tells CAVE in which direction is he oriented, similarly to Hololens GAZE.

One limitation of CAVE system is solving the obstacle of simulating user's movement. This can be resolved by using virtual avatar in the prototype and controlling his movements by using a joystick. User can direct the avatar through-ought the space this way and IoT sensors can detect his movements and evaluate his position. Controlling avatar's actions can be done with the use of Bluetooth clicker that is component sold alongside with Hololens. This can be a simple way how to utilise existing resources and minimise the difference between Hololens and CAVE user experience, since it is a device used with Hololense as well.

Biggest advantage of CAVE systems in this case is availability and quick and easy option how to manipulate with IoT sensors, provide sensors that are not available in real life, examine specific situations with the use of tailored inputs for the desired outputs of the system and use this sensor however is needed. If a real life sensor is acquired, it can easily substituted the simulated one and change its mode to sending real data to server.

6. Future Work

The current state of work provides simple interaction among Microsoft HoleLens and various IoT devices using cloud server as a means for communication between them. It can support the control of basic household items, such as switching devices on and off, setting thermostats or interacting with basics devices in the home using 2D windows connected to the internet. In future work, we want to focus on using cave automatic virtual environment (CAVE). This system will have VR as master and SmE as slave. This environment will be a safe way for disabled people how to enhance their learning processes for the use of user interfaces through the prototype of their own house. Costs for creating virtual houses with all IoT devices will be cheaper as prototyping in the real environment. We can also check if all sensors work properly in their range and if they are suitable for disabled person in wheelchair. Algorithms determining what kind of user is using the environment could be future addition to the development of proposed system [5]. Using this kind of detection could be used to identify whether the system is used by a specific person, the category he belongs to such as child or a senior, possibly even if the user is disabled. This information could be used to automatically set up the environment according to the user thats using it.

In case the user is a child, controlling of system could be restricted so that the child cannot access options that would endanger him, or adjust the interface so that it

would be easier to understand adequately to his age. This detection functionality can make the interface safer, easier to use and omit the need for user login.

7. Disabled people and user interface

Disabled people have different types of disabilities, not just physical or sensory disabilities, but also many times cognitive impairments, i.e. learning disabilities or an attention disorder. Therefore, when creating a user interface in virtual reality, it should be suitable for such people and adapt to theirs disabilities. User interfaces are therefore specific according to particular type of disability. Disabilities can be categorized as:

- sensory disability, visual, olfactory, auditory,
- physical disability, limited mobility
- mental disability,
- speech disorders
- learning disorders such as dyslexia, dysgraphia, dyscalculia,
- other types such as senility and different dependencies.

People with disabilities need an adjusted user interface to work with standard technologies. New technologies can help them to gain more information and facilitate their daily work. Most work has been done in the area of user interfaces dedicated for hearing and visually impaired people. Support for physically and cognitively affected is sparser. Research [7] is concerned with bringing virtual reality closer to people in wheelchairs or older people.

8. Conclusion

In this paper we have presented Cloud web-based controlling of IoT devices using device for mixed reality Microsoft Hololens. Microsoft HoloLens can immerse user to mixed reality in a very good way, and user can move in space without any limitations compared to using smartphones. When using proposed approach, we have noticed positive advantages when controlling SMART home with the use of one application that collects data from server that integrates all IoT devices in one space. Our goal is to integrate more devices and create one user interface for easy access to all devices and their functionality. Another goal is to create full-featured simulation in CAVE that could help fast prototyping of users' houses, and to transform them into SMART houses.

Acknowledgements

This work has been supported by the APVV grant no. APVV-16-0202 "Enhancing cognition and motor rehabilitation using mixed reality".

References

[1] B. Sobota, Š. Korečko, L. Jacho, P. Pastornický, M. Hudák and M. Sivý. Virtual-Reality Technologies and Smart Environments in the Process of Disabled People Education. In: ICETA 2017, pages 239-246. Danvers: IEEE, 2017.

[2] S. Mayer and G. Soros. User Interface Beaminng – Seamless Interaction winth Smart Things Using Personal Wearable Computers. In: 11th International Conferrence on Wearable and Implantable Body Sensor Networrks Workshops. pages 46-49, IEEE. 2914

[3] M. Sivý and M. Hudák. Ovládanie výpočtovej techniky pomocou očí a hlavy pre hendikepované osoby. In: Virtuálno-realitné technológie v procese vzdelávania hendikepovaných osôb, pages 44-48. TUKE, 2017

[4] R. Furlan. The future of augmented reality: Hololens-microsoft's ar headset shines despite rough edges. In: resources tools and toys. IEEE Spectrum, 53(6):21– 21, 2016.

[5] D. Grotovský et al. OntoSEC: Cloud Service Recognition. Transactions 10.1, pages 52-65, 2017.

[6] M. Hudák, Š. Korečko, B. Sobota. On architecture and performance of LIRKIS cave system. In: 8th IEEE International Conference on Cognitive Informatics, Poprad, pages 117-121, 2017.

[7] T. Ohshima, R. Shibata, H. Edamoto and N. Tatewaki. Virtual ISU: Locomotion Interface for Immersive VR Experience in Seated Position (1). In: SA '16 SIGGRAPH ASIA 2016 Posters Article No. 18. pages 1-2. 2016

[8] OptiTrack: OptiTrack for VR. 2018. [Online.] https://optitrack.com/motion-capture-virtual-reality/

About Authors...

Martin SIVÝ born in 1993 in Humenné, Slovakia. In 2017 he graduated (MSc) at the Department of Computers and Informatics of the Faculty of Electrical Engineering and Informatics at Technical University of Košice. He defended his master's thesis in the field of Informatics. Currently, he is a PhD student in the same department. His research is focused on virtual reality and smart user interfaces.

Marian HUDÁK born in 1992 in Košice, Slovakia. In 2017 he graduated (MSc) at the Department of Computers and Informatics of the Faculty of Electrical Engineering and Informatics at Technical University of Košice. He defended his master's thesis in the field of Informatics. Currently, he is a PhD student in the same department. His research is focused on virtual reality, gesture recognition, and collaborative immersive technologies.

CNN based near-duplicate image detection used for email spam detection

Vít LISTÍK[1]

[1]Dept. of Cybernetics, Faculty of Electrical Engineering,
Czech Technical University in Prague, Technická 2, 166 27 Praha, Czech Republic

listivit@fel.cvut.cz

Abstract. *Near-duplicate image detection is used for searching similar images but may also be used for spam detection. Near-duplicate images may be blacklisted as image spam or number of occurrences of the nearly same image may be calculated. We propose using features extracted from the convolutional neural network for the detection task. We also propose how to convert the vector to a much smaller hash and we compare those methods to previously used methods. We did test our solution for 1000 publicly available images. We did 5 alternations on each image from the dataset (blur, noise, crop, rotate, brightness). Then we tested all the algorithms for searching if the original image is the closest one to the altered images. Our proposed solution performed significantly better for this task. Other algorithms were having a lot of false positives for cropped and rotated images.*

Keywords

Image, Near-duplicate detection, Spam, ResNet, Convolutional Neural Network (CNN)

1. Introduction

In this work, we want to tackle the problem of email spam. Image spam in particular [12]. Spam emails are defined as unsolicited messages which are usually sent in bulks. Image spam refers to images contained in spam messages. This technique was first used to hide from the anti-spam filter (which was not doing image analysis). Nowadays images are very common part of email communication, therefore, they are used for anti-spam with text and other information which may be extracted from emails.

Duplicate detection is an essential part of the anti-spam system because spam messages are sent in bulks. But the messages may contain slight alternations, making exact duplicate detection inefficient. That is why near-duplicate detection is used [16]. Near-duplicates may be detected based on previously gathered samples (blacklist) or they may be transformed to different representation (usually simpler) by which they are clustered and the number of their occurrences over some fixed time window is monitored.

2. State of the art

Near-duplicate image, detection is used for image search, copyright enforcement or spam detection for websites or emails [15]. Image duplicates may be detected by cryptographic hashing the same as for any other data. Near-duplicate detection is much more complicated because similar objects should have same hash or should be nearby some given metric, for which the cryptographic hash is totally unsuitable because by its definition a small variance in input should cause a huge difference in the output. Near-duplicate image detection is usually based on image analysis (like histograms or wavelengths analysis) or locality-sensitive hashing or scale-invariant feature transform (SIFT) [3, 6, 4]. Sift of often superior to the other techniques because it is using higher level features extracted from the image, which on the other hand may affect performance [11].

3. Methods

We are using several image hashing methods. Reference methods are implemented in the ImageHash library [1] and our proposed method is based on convolutional neural network (CNN) and is described in the following chapters.

3.1. Average hashing

Average hashing is the simplest algorithm from Imagehash library and it is also the quickest one [2]. This algorithm is similarly as others based on the idea of reducing high frequencies which carry detail and keeping only low frequencies. The algorithm is the following:

1. Reduce image size (Bilinear) to 8x8 (64) pixels
2. Covert to grayscale

[1] https://github.com/JohannesBuchner/imagehash

3. Compute average color of the whole image
4. Bits are computed based on pixel values which are thresholded by the average to 0 and 1 (64-bit integer)

This algorithm should be immune to scaling and brightness and contrast changes.

3.2. Perceptual hash

Perceptual hash or pHash is described in [2]. This algorithm is based on discrete cosine transform (DCT):

1. Optional size and color reduction (because of speed)
2. Compute the DCT which transform the image to wave spectrum (used in JPEG)
3. Reduce high frequencies from the image
4. Compute average value and do the thresholding (same as for average hashing)

3.3. Difference hash

Difference hash or dHash is described in [1]. This algorithm is based on gradient direction:

1. Size and color reduction (same as for average hashing)
2. Compute relative gradient direction (difference of adjacent pixels)
3. Bits are computed based on the brightness of the neighbor

3.4. Wavelet hash

Wavelet hash or whash is based on Haar wavelet which is based on Fourier analysis. This approach does convert the image to wave spectrum in which it is possible to reduce high frequencies directly.

3.5. Convolutional neural networks

Convolutional Neural networks (CNN) are a special case of artificial neural networks. Those networks are composed of neurons which are not in linear layers but they form filters. Those filters are trained for a specific task. This approach is commonly used for image processing tasks [9]. CNNs are used because of the ability to compress the information from multi-dimensional data like images and because of their high performance. Those networks may be used thanks to raising the computing power of GPUs. Thanks to them the state of the art results for image classification and object detection changed significantly [13]. CNNs trained for classification task on millions of images are great feature extractors also for other tasks. The image features are extracted from the previous to last network layer and are often called semantic feature vector.

3.6. Resnet

One of the examples of those deep CNNs is ResNet [5]. This network architecture won the ImageNet classification challenge in 2015 [14]. The architecture is using residual connections which made training of network this deep possible.

3.7. CNN hash

We propose using a semantic vector extracted from the CNN as a hash [7]. This method should be invariant to alternations because the network is trained to compress important visual information to the extracted feature vector.

We did also propose how to create a universal representation from this vector. The binarization is inspired by the average hash. We did the binarization based on fixed threshold 0.5 and also based on the mean. The extracted vector has length 2048, therefore it is space inefficient and keeps too much information. We also compress this binarized vector with a technique similar to max pooling. We took windows of size 2, 4, 8, 16, 32, 64 and searched for at least one 1 (max reduction) and looked if the average is above 0.5 (average reduction).

3.8. Distance metrics

We need some way to do define the distance of two image representations (hashes). There are two metrics used in this work.

Hamming distance - is a metric used for a distance of strings. This metric computes the number of changes needed for inputs of the same length to make them equal. This metric is used in ImageHash library for comparison of the bits of the 64-bit integer.

Cosine distance - Is used for real-valued high dimensional vectors. This metric ignores the magnitude of the vectors. The computation of cosine distance is described in Eq. 1.

$$\cos(\Theta) = \frac{A.B}{\|A\| \, \|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}} \quad (1)$$

3.9. Implementation

Our implementation may be found at Github platform[2] as an open-source. Implementation is in Python language and is runnable via Docker compose.

4. Experimental results

We performed several experiments on publicly available images and their alternations which show the performance of described hashing methods.

4.1. Dataset

We used 1000 first images from OpenImages database [10, 8]. The dataset consists of photos which may be used for classification or object recognition. We used this dataset because it is very general and will be similar to the email traffic which would be much harder to obtain because of the email private nature.

4.2. Image alternations

We did alternate the images for testing the robustness of the hashing approach. The alterations were the following:

- *Blur* - Gaussian blur of the image pixels with radius 2
- *Crop* - Random crop of the image not smaller then 40% of the original
- *Brightness* - Making the image lighter or darker up to 80%
- *Rotate* - Image rotation up to 30 deg
- *Noise* - Adding Gaussian noise to the image

4.3. Similarity detection

We did test if the alternations of the image are the closest ones to the original image. The images were sorted based on the distance metric and if the alternation was found after some other images the score was lowered proportionally to the number of false positive images (0-1) shown in Tab. 1.

We also evaluated each alternation separately. The ordering for the CNN is unique, the exact ordering for other hashing methods may be uncertain, therefore the alternated images "compete" with each other for CNN that may share the same spot for the other methods. The results are shown in Tab. 2

[2] https://github.com/tivvit/image-duplicate-detection-eval

method	score
cnn	0.98
average hash	0.10
dHash	0.04
pHash	0.03
wHash	0.09

Tab. 1. Score in range 0-1 for searching image alternations

method	Blur	Crop	Bright	Rotate	Noise
cnn	2.39	4.44	1.65	2.97	4.03
avg hash	1.05	802.00	2.68	350.36	1.63
dHash	1.20	1843.44	3.00	559.65	3.52
pHash	1.11	1761.22	2.39	756.63	1.69
wHash	1.07	755.39	1.87	310.55	1.39

Tab. 2. Average positions of the alternations for the hashing methods.

4.4. Hash representation

We tested if the altered image hash representation is an exact match with the original image representation. The results are shown in Tab 3.

method	Match
dHash	33.2%
pHash	42.5%
average hash	44.5%
wHash	50.7%
ab8	0.4%
aba16	0.1%
ab16	2.5%
mb16	5.5%
mba16	5.9%
ab32	6.3%
aba32	7.0%
ab64	10.1%
mb32	12.3%
aba64	12.5%
mba32	13.0%
mb64	13.1%
mba64	13.1%

Tab. 3. Results for number of exactly matching hashes. Where ab8 means average reduced windows of size 8 for threshold binarization (ba - average binarization, m - max reduction).

5. Conclusions and future work

We show the performance of near duplicate image detection algorithms on 1000 images. Each of the images was altered with 5 operations (blur, noise, crop, rotate, brightness). We did search the closest images to the original image and shown that proposed CNN solution with score 0.98 (max

1) is significantly better for that task then simpler similarity detection algorithms which scored best 0.1. The feature vector extracted from CNN was performing almost the same for all alternations. Other hashing algorithms were affected by rotation and mostly by crop. We did also test the exact match of the computed hash from the altered image to the hash computed from the original image. Our simple method for binarization and reduction the feature vector achieved up to 13% match while hashing algorithms achieved up to 50%.

Our proposed solution of an image near-duplicate detection using CNN extracted feature vector is able to perform well for known samples. When we want to use the image hash as an identifier it is better to use wHash. Both methods are usable for anti-spam detection. The hashing solution may be used for counting the number of occurrences of the images in the traffic and the solution using CNN feature vector may be used for the exact match. The exact match may be used for blacklisting or a tree structure with image representations may be used for creating image buckets for counting occurrences.

Acknowledgements

The research described in the paper was supervised by Prof. V. Hlaváč and J. Šedivý CSc. CIIRC in Prague.

References

[1] Kind of like that - the hacker factor blog.

[2] Looks like it - the hacker factor blog.

[3] Ondrej Chum, James Philbin, Andrew Zisserman, et al. Near duplicate image detection: min-hash and tf-idf weighting. In *BMVC*, volume 810, pages 812–815, 2008.

[4] Wei Dong, Zhe Wang, Moses Charikar, and Kai Li. High-confidence near-duplicate image detection. In *Proceedings of the 2nd acm international conference on multimedia retrieval*, page 1. ACM, 2012.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[6] Yan Ke, Rahul Sukthankar, Larry Huston, Yan Ke, and Rahul Sukthankar. Efficient near-duplicate detection and sub-image retrieval. In *ACM multimedia*, volume 4, page 5. Citeseer, 2004.

[7] Giorgos Kordopatis-Zilos, Symeon Papadopoulos, Ioannis Patras, and Yiannis Kompatsiaris. Near-duplicate video retrieval by aggregating intermediate cnn layers. In *International conference on multimedia modeling*, pages 251–263. Springer, 2017.

[8] Ivan Krasin, Tom Duerig, Neil Alldrin, Vittorio Ferrari, Sami Abu-El-Haija, Alina Kuznetsova, Hassan Rom, Jasper Uijlings, Stefan Popov, Shahab Kamali, Matteo Malloci, Jordi Pont-Tuset, Andreas Veit, Serge Belongie, Victor Gomes, Abhinav Gupta, Chen Sun, Gal Chechik, David Cai, Zheyun Feng, Dhyanesh Narayanan, and Kevin Murphy. Openimages: A public dataset for large-scale multi-label and multi-class image classification. *Dataset available from https://storage.googleapis.com/openimages/web/index.html*, 2017.

[9] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[10] Alina Kuznetsova, Hassan Rom, Neil Alldrin, Jasper Uijlings, Ivan Krasin, Jordi Pont-Tuset, Shahab Kamali, Stefan Popov, Matteo Malloci, Tom Duerig, and Vittorio Ferrari. The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale. *arXiv:1811.00982*, 2018.

[11] David G Lowe et al. Object recognition from local scale-invariant features. In *iccv*, volume 99, pages 1150–1157, 1999.

[12] Bhaskar Mehta, Saurabh Nangia, Manish Gupta, and Wolfgang Nejdl. Detecting image spam using visual features and near duplicate detection. In *Proceedings of the 17th international conference on World Wide Web*, pages 497–506. ACM, 2008.

[13] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015.

[14] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, 2015.

[15] Ramarathnam Venkatesan, S-M Koon, Mariusz H Jakubowski, and Pierre Moulin. Robust image hashing. In *Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101)*, volume 3, pages 664–666. IEEE, 2000.

[16] Zhe Wang, William K Josephson, Qin Lv, Moses Charikar, and Kai Li. Filtering image spam with near-duplicate detection. In *CEAS*, 2007.

Using sentiment detection on tweets to form opinions

Jakub Konrád[1]

[1]Dept. of Cybernetics, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

konrajak@fel.cvut.cz

Abstract. *In this paper we explore the possibility of using social media to aggregate knowledge about previously unknown subject and use sentiment analysis to gauge the overall opinion on a certain topic on social media.*

We train different sentiment classifier models and leverage them together with TwitterAPI to form a sentiment score on a presented topic that serves as a basis for the system's opinion.

Keywords

sentiment analysis, opinion, BERT, twitter

1. Introduction

Formulation of opinion about previously unencountered topics is a difficult but crucial task in conversational artificial intelligence. When engaging in conversation with the user, we are often presented with topics we have not met before. Here the system needs to quickly form an opinion about the presented topic.

In this paper, we try to leverage social media networks, specifically Twitter to gain a general sense of the sentiment connected to the new topic user presented the system. We use sentiment classification to achieve this. We have selected two datasets as sources for our training data.

Firstly the Large Movie Review Dataset (aclIMDB) presented in [1]. The dataset serves for binary sentiment classification. It provides a set of 25,000 highly polar movie reviews for training and 25,000 for testing. This dataset is well known and often used as a base for sentiment classification tasks. Secondly we utilize the Sentiment 140 dataset presented in [2]. The dataset contains tweets obtained via Twitter API. Tweets were labeled as positive or negative based on emoticons used in any given tweet.

We train sentiment classifiers using these datasets firstly using a simple recurrent neural network, and secondly, we use the BERT [3] text classification task on the same datasets. We compare the obtained results and evaluate the differences.

2. Sentence sentiment classification approaches

We use two approaches for sentence entity detection, a bidirectional GRU[4] neural network for which we train a model from scratch, and a BERT [3] pretrained model that we fine-tune for sentiment classification on our data. We train the models twice on different sets on data, the aclIMDB [1] dataset and once on the Sentiment140 dataset [2].

2.1. GRU sentiment

Firstly the text is cleaned of all non-word strings and then tokenized. Individual tokens are converted to their embedding representation. We utilize the standard GLoVe [5] embeddings of the length 50. These embeddings are then fed as input into a bidirectional (GRU) recurrent neural network layer. We then obtain the final sentiment of the text via a dense layer.

With the described setup we have achieved the accuracy 88.2% on the validation set for the aclIMDB dataset and the accuracy 83.1% for the sentiment140 dataset.

2.2. Sentiment detection using BERT

We use BERT [3] as the second method of sentiment classification. BERT (Bidirectional Encoder Representations from Transformers) is a state of the art natural language representation model designed to pretrain deep bidirectional representations. The main advantage of the model is that it is possible to fine-tune the pretrained representations for a variety of NLP tasks [3]. In our work, we use BERT for sentiment classification purposes.

We achieve an accuracy of 89.4% on the validation set for the aclIMDB dataset and accuracy of 85.6% for the sentiment140 dataset.

3. Forming opinions using sentiment

Now we need to describe the opinion formulation we use in our system. We use entity detection to find relevant

topics, and then we scrape twitter via TwitterAPI[1] for tweets containing the detected entity. We then classify all the tweets using one of our trained sentiment classifiers and compute a mean sentiment score from all the tweets. The Sentiment score takes a value between 0 (negative) and 1 (positive) that we consider the system's opinion on the entity.

4. Evaluation

4.1. Comparison of GRU and BERT models

Firstly let us discuss the different types of classifiers used for this experiment and their results. On both datasets the BERT model achieved slightly higher accuracy than the GRU model (Table 1). The relative improvement amounts to 1.36% for aclIMDB and 3.0% for Sentiment140.

Dataset	aclIMDB	Sentiment140
GRU Model	88.2%	83.1%
BERT	89.4%	85.6%

Tab. 1. Comparison of the performance of GRU model and BERT on both datasets.

4.2. Comparison of results based on training data

Secondly, we compared the final sentiment scores obtained using models trained on the aclIMDB and Sentiment140 (See figure 1). During our testing, we have noticed that models with different training data often present significantly different sentiment score. Specifically, the models trained on the Sentiment140 dataset provided a sentiment score that is on average 32.7% more positive then the aclIMDB models, with the mean of the aclIMDB sentiment score sitting at 0.486 and mean of the Sentiment140 models at 0.722.

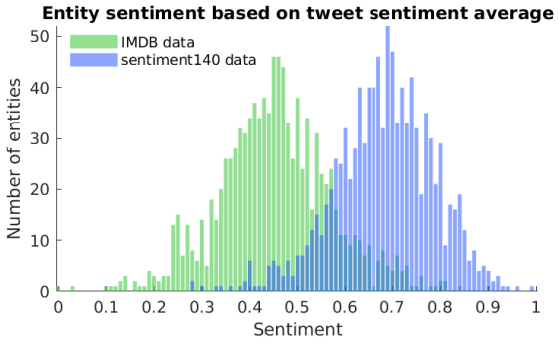


Fig. 1. Histogram of detected entity sentiments for models trained on aclIMDB and sentiment140. Sentiment value 0 translates to the most negative and sentiment value 1 as the most positive.

[1] https://developer.twitter.com/en/docs/api-reference-index

5. Conclusion

We have proposed a simple method for the formulation of opinion about previously unencountered topics during a conversation. We have tested this approach with a simple GRU model and BERT fine-tuned model with two sources of training data.

We have shown that using BERT we can achieve slightly better results than using the GRU model. However, we need to take into account the complexity of both systems. While the better performance of BERT might be significant it is better suited for offline applications where its high hardware requirements do not pose such a limitation.

We have also compared models trained on Sentiment140 and on aclIMDB. While Sentiment140 is a dataset composed of actual tweets we have found the obtained sentiment scores carry huge positive bias that often leads to results that were not only worrying but also dangerous if they were incorporated into a system interacting with real users (e.g., Hitler 0.64 vs. 0.28, murder 0.38 vs. 0.24). Due to this, we have elected to use only the aclIMDB models for our system.

Acknowledgements

The research described in the paper was supervised by Ing. Jan Šedivý, CSc., CIIRC CTU in Prague and it was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS18/103/OHK3/1T/37.

References

[1] Maas, A. L.; Daly, R. E.; Pham, P. T.; et al. Learning Word Vectors for Sentiment Analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Portland, Oregon, USA: Association for Computational Linguistics, June 2011, pp. 142–150. Available from: http://www.aclweb.org/anthology/P11-1015

[2] Go, A.; Bhayani, R.; Huang, L. Twitter Sentiment Classification using Distant Supervision. *Processing*, 2009: pp. 1–6. Available from: http://www.stanford.edu/ alecmgo/papers/TwitterDistantSupervision09.pdf

[3] Devlin, J.; Chang, M.-W.; Lee, K.; et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[4] Cho, K.; van Merrienboer, B.; Gülçehre, Ç.; et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. *CoRR*, volume abs/1406.1078, 2014, 1406.1078. Available from: http://arxiv.org/abs/1406.1078

[5] Pennington, J.; Socher, R.; Manning, C. D. GloVe: Global Vectors for Word Representation. In *Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1532–1543. Available from: http://www.aclweb.org/anthology/D14-1162

About Authors...

Jakub KONRÁD graduated with master's degree in Jan 2018 with a thesis on Unstructured Text Comprehension and Question Answering. He co-created Alquist Dialogue Manager, a framework for bot creation. Since 2017 he has been working on the Alquist bot for the first Amazon Alexa prize. He continues to work on conversational AI during his PhD.

Automatic Test of Visual Feature Detectors and Descriptors for Visual Odometry

Tomáš Pivoňka[1]

[1]Dept. of Cybernetics FEE and Intelligent and Mobile Robotics CIIRC,
Czech Technical University, Jugoslávských partyzánů 1580/3, 160 00 Praha, Czech Republic

pivontom@fel.cvut.cz

Abstract. *Performance of feature-based visual odometry highly depends on a used method for feature detection and description. It is not influenced only by a type of a method but also by its inner parameters' setting. Therefore the automatic test of feature detectors and descriptors is presented in this paper. It allows comparing various methods with different inner parameters. The test evaluates features' properties on a small dataset of images. It was used for finding the most suitable method for visual odometry in the camera-based scanner of car undercarriages.*

Keywords

Visual features, Computer vision

1. Introduction

Visual features are key parts of many computer vision applications. They represent local parts of images marked with descriptors, which allow matching the same features in different images. These correspondences are used in feature-based visual odometry methods to compute rotation and translation between images captured from different views. This approach fails, if not enough correspondences are found, or a lot of them are incorrect. Therefore the quality of a method for feature detection and description highly influences precision and robustness of visual odometry.

This work is a part of a project for developing a camera based car undercarriage scanner. This scanner serves to a security check of cars, and it returns a 3D model of an undercarriage. The scanner consists of several stereo cameras which return particular 3D models. They are stitched together to create the model of the whole undercarriage. Transformations between models are computed by visual odometry. The visual odometry method is described in [1].

Images of car undercarriages differ from scenes usually used in visual odometry, and they contain a lot of low textured surfaces. A motion of a car is constrained, and the most dominant motion is a forward car movement. Other rotations and translations are caused by a slight turning of a car or by vibrations.

Basic principles of the tested methods for feature detection and description and its inner parameters in OpenCV are presented in Sect. 2. The automatic test was used to find an optimal method for feature detection and description. System performance does not depend only on a type of a used method, but an inner parameters' setting highly influences it. The automatic test searches also for the best setting of inner parameters and it is described in Sect. 3. The results are presented in Sect 4.

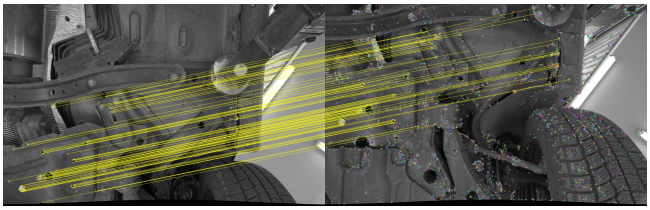


Fig. 1. Undercarriage images with detected and correctly matched features.

2. Visual Features

Visual features are parts of images well distinguishable from their neighbourhood. They can differ in intensity, colour or texture. There are many of various approaches to feature detection and description. In this work, there are used standard methods from the OpenCV library [3]. OpenCV is an open source library, and it contains the majority of key algorithms used in computer vision. Properties of detected features highly depend on inner parameters' setting of the used method. Basic principles and inner parameters of used methods are presented below in this section.

The Harris corner detector [4] is an improvement of the Moravec detector, which is one of the oldest corner detectors. The main idea is that shifts of a local window in all directions significantly change its appearance if the window contains a corner. The Harris detector uses a Gaussian window, in which it computes matrix M from approximations of partial derivatives. Matrix M is called a local autocorrelation

function, and it serves to calculate parameter R. A window contains a corner if its parameter R is higher than an assigned threshold.

The Shi-Tomasi detector [5] is very similar, but its parameter R is a minimal eigenvalue of matrix M.

Harris detector has several inner parameters: the size of a window for feature detection, the size of a Sobel filter window for computing partial derivatives, parameter k used for calculating R and a distance for removing close corners. The Shi-Tomasi detector has the same parameters as the Harris detector without parameter k.

FAST [6] is a corner detector based on searching pixels on a circle around tested points. A feature is detected if there is a continuous sequence of points on a circle which are brighter or darker than the tested point. Its inner parameters are a threshold, which states how much brighter or darker pixels have to be, using of non-maximum suppression for filtering close features, the size of a circle and a minimum length of a sequence. In OpenCV implementation, the size of a circle and the minimum length are related and cannot be chosen independently.

The BRIEF descriptor [7] is a binary descriptor. Binary descriptors are smaller than numerical descriptors, and Hamming distance gives a distance between descriptors. A descriptor is a vector of binary values, and each value is a comparison of intensity between two selected points in feature neighbourhood. There are various approaches to selection of point pairs - e.g. random, Gaussian or uniform distribution of points.

Methods for creating descriptors have own inner parameters too. They are the size of a descriptor (number of bits) and invariance to rotation, but rotation has to be computed during detection of feature, and it is not detected in a method for feature description.

SIFT [8] incorporates feature detector and descriptor, which is highly invariant to scale and rotation transformations. It is based on a Gaussian filtration of an image with different variance. Filtered images are split to octaves, and each octave contains a set of images (levels) with increasing variance. The features are local maxima in images, which are differences of neighbouring images in octaves (i.e. with different variances). The maximum is searched in neighbouring pixels in the same and adjacent levels. In the next step, edge or weak features are removed. Rotation is determined from a histogram of gradient orientations in a feature point vicinity. In the next step, a point vicinity is split to a matrix of square areas and descriptor is created from histograms of gradients with eight directions in each square.

Inner parameters of SIFT are the number of layers in one octave, a threshold to remove weak features and a threshold to remove edge features (same as in Harris detector), an initial value of variance and the maximal number of returned features.

Among the main disadvantages of SIFT are high computational demands. SURF [9] is a similar method, but it is much faster. A difference of Gaussians is replaced by rectangular filters, which return approximations of partial derivations. Hessian matrix is composed of partial derivatives, and features are searched as a maximum of a determinant in a point neighbourhood. Filtered images also create octaves and layers. Feature orientation is computed by wavelength transform. Wavelength transform with Haar wavelet is also applied to square areas in feature vicinity to create descriptor.

SURF inner parameters are the number of octaves, the number of layers in one octave, a threshold for a determinant of Hessian matrix, the size of a descriptor and using computation of the feature orientation. Orientation does not have to be computed, but it improves robustness to rotation changes.

The last tested method is ORB [10] (Oriented FAST and Rotated BRIEF). Detection is based on the FAST detector, but it is performed on differently scaled images to ensure invariance to scale changes. Features are evaluated by parameter R from the Harris detector, which enables to select the required number of the best features. Orientation is computed from a position of the centre of intensity. The orientation is used to rotate pairs of points to create BRIEF descriptor.

The ORB method has many inner parameters: the maximal number of returned features, the number of layers, a scale factor between layers, a FAST detector threshold, using not binary descriptor (computed from 3 or 4 points), and a type of evaluation from Harris or FAST detector.

3. Automatic Test

In the automatic test, features are detected and matched between images from a stereo camera. Use of a stereo camera allows using epipolar geometry for evaluation of the correctness of matches. The test is performed on a small dataset of undercarriage's images, and a method and its parameters with the best results on this type of scene is searched.

For each inner parameter, a set of values is selected, and various combinations of parameters' values are tested on the dataset. The number of values for each parameter is selected considering the time of testing all combinations. Values are chosen to cover a space of possible values uniformly and based on previous experiments.

Each combination of parameters' values is tested on the dataset separately. In every stereo image, features are detected and matched. Images are divided into 12 rectangular regions during detection, which ensures a more uniform distribution of features. Matched features are tested, if the point's distance from an epipolar line, which is computed from the corresponding point in the second image, is smaller than a selected threshold. If this condition is satisfied, a 3D

position of the point is triangulated. An approximate distance from the camera is known, so a distance of correctly matched point lies in a known interval. If both conditions are satisfied, a correspondence is considered to be correct.

The measured parameters for each combinations' test are the number of correctly matched features, the minimal number of correctly matched features in one stereo image, the number of matched features in all images and the maximal computing time for one stereo image. The most important is the second parameter because it ensures to have enough features in all images. Visual odometry usually need-not a lot of features; for example, methods based on solving perspective three point problem can work with only four correct matches.

4. Results

Because a prototype of a scanner was not available, the scanner was simulated by a car on a host and moving cameras under it. Images were captured by Basler acA1920-155uc colour camera with the resolution of 1920x1200px. Four cameras were placed on a special construction, and they created two stereo pairs. Each pair took pictures of one half of the undercarriage. The images were taken in two different distances and with two different exposition times, which simulate various lighting conditions. In the automatic test, images were evaluated separately for each exposition time, because a scanner with lighting can ensure stable lighting conditions. Images are divided into dark and light dataset according to used exposition time. Cameras' distances from undercarriage were approximately 57 cm and 75 cm.

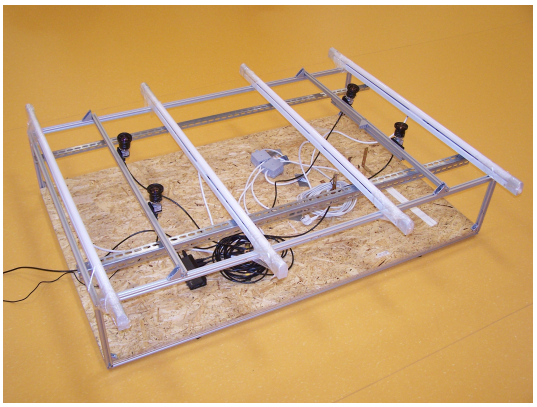


Fig. 2. Construction with cameras for capturing images of undercarriage by shift it under a car on a host.

Feature detectors and descriptors described in Sect. 2 were tested in this work. The tested detectors were Harris, Shi-Tomasi, FAST, SIFT, SURF and ORB. SIFT, SURF and ORB are not only detectors, but they also have a method for feature description. Remaining methods were tested with the BRIEF descriptor. All these methods are implemented in the OpenCV library [3].

	Harris	Shi-Tomasi	FAST	SIFT	SURF	ORB
Min. matches 1 img.	5	4	0	137	49	158
Correct matches	162	154	104	3 748	2 384	4 402
Matches	21 492	17 146	8 464	35 718	24 310	41 756
Time [s]	0.26	0.26	0.072	2.13	0.49	0.38

Tab. 1. Feature detectors' comparison for dark images [1]

	Harris	Shi-Tomasi	FAST	SIFT	SURF	ORB
Min. matches 1 img.	6	5	9	258	356	194
Correct matches	176	190	590	6 162	9 822	5 198
Matches	25 512	24 746	52 566	36 276	57 834	37 570
Time [s]	0.26	0.29	0.68	4.97	0.96	0.36

Tab. 2. Feature detectors' comparison for light images [1]

Results of the automatic test are presented in Tab. 1-3. The presented results come from the best inner parameters setting for each method in the dataset. The most important parameter was a minimal number of matches in one stereo image and the second was the number of correct matches in all images. In some cases, maximal time of detection has to be considered, because it was longer than 1 s. Even though some methods with longer computational time can achieve better results, they cannot be used in our application, because of demands to almost real-time system performance. Tab. 1 shows performance of methods in dark images. The testing on the light dataset is summarized in Tab. 2. The results for the whole dataset (dark+light) are presented in Tab. 3.

Harris and Shi-Tomasi detectors have very similar results in this application. They detect enough number of features, but they have not got sufficiently strong descriptors, and most of the matches are incorrect. For example, the Harris detector found 25 512 matches in the dark dataset, but only 176 of them were assessed to be correct.

	Harris	Shi-Tomasi	FAST	SIFT	SURF	ORB
Min. matches 1 img.	4	4	0	137	49	158
Correct matches	150	161	371	4 737	6 103	4 899
Matches	23 787	24 892	29 950	35 701	41 072	39 773
Time [s]	0.28	0.29	0.77	2.13	0.96	0.39
Combinations	1500	300	144	288	384	800

Tab. 3. Feature detectors' comparison for all images [1]

The FAST detector works better for light images, but it still has an insufficient number of correct matches. In addition, only 1.1 % of matches were correct.

SIFT, SURF and ORB reached much better results. SIFT has the second best results in all datasets, but none of the SIFT parameters' combinations has maximal computing time shorter than 1.2 s; therefore it cannot be used in this application. Only methods with a time below 1 s were searched because the scanner has to display results as soon as possible.

If the computation time were not limited, SURF would exhibit the best results. But only parameters' combinations with time under 1 s were considered. On the other hand, even the best results do not overcome ORB significantly, because a minimal number of correct matches in one image is 191 and its maximal time for one image is 9.2 s. SURF has the best results in light images. SURF is a suitable method for our application; nevertheless, its usage is limited by its licence.

ORB has the best results in the whole dataset, and it was able to find at least 158 correct matches in each image. The ratio of correct matches in all matches is 12 %. ORB also has the best performance on the dark dataset. In addition, ORB is faster than SURF. Another advantage is that ORB is license free. Finally, the ORB method was selected for the aforementioned visual odometry system.

5. Conclusion

The automatic test allowed comparing different methods for feature detection and description and finding their optimal parameters. Detectors and descriptors implemented in OpenCV were used. None of the corner detectors with BRIEF descriptor was able to find enough number of correct matches in all images. SIFT cannot be used because of its long computational time. ORB and SURF satisfied all requirements on the method. The advantages of ORB were its speed, better results on the whole dataset and its free licence in comparison with SURF.

The ORB method with its optimal parameters was successfully used in the visual odometry system. Because images came only from one car, it is necessary to verify presented results in a larger dataset.

Acknowledgements

The work described in this paper was funded by the Ministry of Interior of the Czech Republic under project VI2VS/461. It was supervised by Ing. Libor Přeučil, CSc. and investigated in cooperation with Ing. Karel Košnar, Ph.D. and Mgr. Martin Dörfler.

References

[1] PIVOŇKA, T., KOŠNAR, K., DÖRFLER, M., PŘEUČIL, L., 2019. Visual Odometry for Vehicles' Undercarriage 3D Modelling. *In: Mazal J. (eds) Modelling and Simulation for Autonomous Systems*. MESAS 2018. Lecture Notes in Computer Science, vol 11472. Springer, Cham

[2] SCARAMUZZA, D., FRAUNDORFER, F., YU, W. Visual Odometry [Tutorial]. *International Journal of Control, Automation and Systems*. 2011, 18(4), 80-92. DOI: 10.1109/MRA.2011.943233. ISSN 1070-9932.

[3] BRADSKI, G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*. 2000.

[4] HARRIS, C., STEPHENS, M. A Combined Corner and Edge Detector. [online]. 1988. Retrieved from: http://www.bmva.org/bmvc/1988/avc-88-023.pdf

[5] SHI, J., TOMASI, C. Good features to track. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94* [online]. 1994. P.593-600. DOI: 10.1109/CVPR.1994.323794. Retrieved from: http://ieeexplore.ieee.org/document/323794/

[6] ROSTEN, E., DRUMMOND, T. Machine Learning for High-Speed Corner Detection. *Computer Vision — ECCV 2006* [online]. 2006. P.430-443. DOI: 10.1007/11744023_34. Retrieved from: http://link.springer.com/10.1007/11744023_34

[7] CALONDER, M., LEPETIT, V., STRECHA, Ch., FUA, Pascal. BRIEF: Binary Robust Independent Elementary Features. *Computer Vision — ECCV 2010* [online]. 2010. P.778-792. DOI: 10.1007/978-3-642-15561-1_56. Retrieved from: http://link.springer.com/10.1007/978-3-642-15561-1_56

[8] LOWE, D. G. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* [online]. 2004. Vol.60, no.2p.91-110. DOI: 10.1023/B:VISI.0000029664.99615.94. Retrieved from: http://link.springer.com/10.1023/B:VISI.0000029664.99615.94

[9] BAY, H., TUYTELAARS, T., VAN GOOL, L. SURF: Speeded Up Robust Features. *Computer Vision — ECCV 2006* [online]. 2006. P.404-417. DOI:10.1007/11744023_32. Retrieved from: http://link.springer.com/10.1007/11744023_32

[10] RUBLEE, E., RABAUD, V., KONOLIGE, K., BRADSKI, G. ORB: An efficient alternative to SIFT or SURF. *2011 International Conference on Computer Vision* [online]. 2011. P.2564-2571. DOI:10.1109/ICCV.2011.6126544. Retrieved from: http://ieeexplore.ieee.org/document/6126544/

About Authors...

Tomáš PIVOŇKA was born in 1993 in Prague. He graduated from FEE CTU in 2018 with diploma thesis Visual Odometry for Dynamic Image Reconstruction. He is a PhD student at the Dept. of Cybernetics, CTU and a member of the Intelligent and Mobile Robotics group of CIIRC CTU.

Optimal evaluation of risky opportunities

Ing. Matej Uhrin [1]

[1]Dept. of Computer Science, Czech Technical University, Karlovo namesti 13, 121 35, Prague, Czech Republic

uhrinmat@fel.cvut.cz

Abstract. *We investigate the problem of optimal wealth allocation over predictive sports market's opportunities. We analyze the problem across diverse input variables, target utility measures, and the notion of optimality itself. Namely, we focus on the provably optimal long run policy, typically referred to as the Kelly Criterion, and Modern Portfolio Theory based approaches leveraging utility theory. From the joint perspective of decision theory, we discuss their unique properties and assumptions. Finally we present an experiment on a dataset from relevant domain of football betting.*

Keywords

portfolio optimization, machine learning, sports analytics

1. Introduction

The main goal of all the methods discussed here is the optimal decision when presented with single or multiple risky opportunities. A dynamic is always defined for the presented opportunities. Dynamic specifies progression of our wealth in time, additionally it specifies if & how the presented opportunities will be repeated. The methods presented in this paper always assume multiplicative dynamic, i.e. wealth is reinvested after every single time step on the same set of opportunities, (We could also not reinvest our wealth and work with unit based investments which we will refer to as additive dynamic). Our wealth under multiplicative dynamic is defined as follows.

$$x(t) = r(t) \cdot x(t - \delta t) \tag{1}$$

- $x(t)$ is wealth in time t.
- δt is a regular time interval, e.g. 1(week).
- $r(t)$ is a return/payoff function. e.g. a favourable game of coin toss, if result is heads our wealth grows by 50% and if it is tails our wealth shrinks by 40%. Then the return function $r(t)$ is defined as follows:

$$r(t) = \begin{cases} 1.5 & \text{with probability } 1/2 \\ 0.6 & \text{with probability } 1/2 \end{cases} \tag{2}$$

Clearly the presented coin toss game is favourable. Assume that if we accept, we are obliged to play this game repeatedly every single day. Is it still a good deal? The following two quantities are relevant to our decision, expectation value, (or ensemble average) $\mathbb{E}[r]$ and the time average $\bar{r}$ as defined in (Peters et al., 2016).

$$\mathbb{E}[r] = \frac{1}{2} \cdot 1.5 + \frac{1}{2} \cdot 0.6 = 1.05 \tag{3}$$

$$\bar{r} = (1.5 \cdot 0.6)^{\frac{1}{2}} \approx 0.95 \tag{4}$$

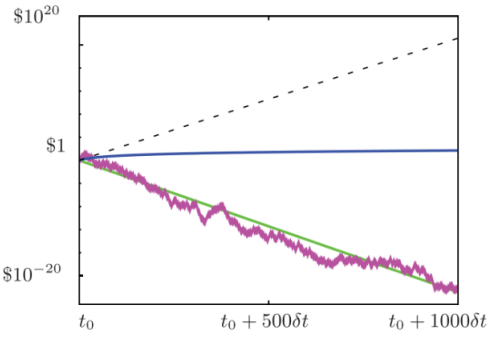


Fig. 1. The magenta line represents median wealth trajectory in 1000 time steps of coin toss game. The dashed line is the expectation value of our game.

Wealth growth under multiplicative dynamic is non-ergodic process, hence the expectation value does not indicate what happens to a single player in the long run.

2. Modern Portfolio Theory

We will denote the vector of wealth fractions as $\boldsymbol{b}$ and refer to it as portfolio.

$$\boldsymbol{b} = \left[b_1, b_2, ..., b_c\right] \tag{5}$$

Where b_c stands for "cash" fraction of wealth, i.e. a fraction we leave on the side, e.g. in a bank or inside a sock. The idea behind Modern Portfolio Theory, (MPT) is that portfolio $\boldsymbol{b_1}$ is superior to $\boldsymbol{b_2}$ if the expected gain $\mathbb{E}[\boldsymbol{b}]$ is at least as great.

$$\mathbb{E}[\boldsymbol{b_1}] \geq \mathbb{E}[\boldsymbol{b_2}] \tag{6}$$

and the risk, here general risk measure denoted r is no greater, (Markowitz, 1952).

$$r(\boldsymbol{b_1}) \leq r(\boldsymbol{b_2}) \tag{7}$$

This creates a partial ordering on the set of all available portfolios. Taking the portfolios that no portfolio is superior gives us the set of efficient portfolios Θ. MPT can be expressed as the following maximization problem:

$$\begin{aligned} \underset{\boldsymbol{b}}{\text{maximize}} \quad & \boldsymbol{\mu}^T \boldsymbol{b} - \gamma \boldsymbol{b}^T \Sigma \boldsymbol{b} \\ \text{subject to} \quad & \sum_{i=1}^{K} b_i = 1.0,\ b_i \geq 0 \end{aligned}$$

where $\boldsymbol{b}$ is vector of wealth fractions, (portfolio), γ is risk aversion parameter and $\boldsymbol{\mu}$ is the expected gains vector of offered opportunities.

In layman terms we maximize the following:

$$\mathbb{E}[gain] - \gamma \cdot risk \tag{8}$$

In the most general set up risk is defined as variance Σ.
This approach is criticised for two main reasons.

1. The growth of wealth is non-ergodic process, (Peters et al., 2016). In other words, the expected gain does not tell us what will really happen in the long run, hence maximizing it does not produce truly long term optimal strategy, which we showcased in the coin toss example.
2. The definition of risk as variance is often disputed. In many domains, risk is not easy to define.

3. Growth optimal approach

A different approach is to first transform the process into ergodic process and then take the expectation value. This approach is famously known as the "Kelly Criterion", (Kelly Jr, 2011) and many other names. The algorithm is as follows:

1. Specify the wealth dynamic, i.e. the relationship between $\delta x(t)$, $x(t)$ and t.
2. Find the ergodicity transformation, $v(x)$, of the wealth x whose increments are instances of a (time-independent) random variable under specified dynamic.
3. Determine the time-average growth rates, either by taking the long-time limit of the growth rates or by invoking the ergodic property and taking their ensemble averages.
4. Choose portfolio $\boldsymbol{b}$ with the largest time-average growth rate.

For a multiplicative dynamic, the correct ergodicity transformation is a logarithm: $v(x) = \log(x)$, hence the growth optimal portfolio can be found by solving the following optimization problem.

$$\begin{aligned} \underset{\boldsymbol{b}}{\text{maximize}} \quad & \mathbb{E}[\log(\boldsymbol{R} \cdot \boldsymbol{b})] \\ \text{subject to} \quad & \sum_{i=1}^{K} b_i = 1.0,\ b_i \geq 0 \end{aligned}$$

- K probabilistic outcomes $p_1, p_2, ..., p_K$.
- n opportunities, $n-1$ risky, 1 risk-less. $\boldsymbol{a_1}, ..., \boldsymbol{c}$.

$$\boldsymbol{p} = \begin{bmatrix} p_1 & p_2 & ... & p_K \end{bmatrix} \quad \boldsymbol{a_i} = \begin{bmatrix} r_{1,i} \\ r_{2,i} \\ ... \\ r_{K,i} \end{bmatrix} \quad \boldsymbol{c} = \begin{bmatrix} 1 \\ 1 \\ ... \\ 1 \end{bmatrix} \tag{9}$$

Cash asset can have a different payoff if money can be **risk-free invested** elsewhere. (e.g. bank acc interest rate).

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{a_1} & \boldsymbol{a_2} & ... & \boldsymbol{a_{n-1}} & \boldsymbol{c} \end{bmatrix} \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ ... \\ b_{n-1} \\ b_c \end{bmatrix} \tag{10}$$

where $\boldsymbol{R}$ is a return matrix and $\boldsymbol{b}$ is a portfolio vector.

The calculated portfolio is growth optimal under the following assumptions: The long run consists of approximately similar opportunities and true probability distribution of the outcomes is known to the investor.

4. Experiment

We showcase the growth optimal approach on a dataset of football matches. We have three outcomes per game and we assume 10 parallel games happening simultaneously in each time step. The probabilities are modelled separately using neural network model. The true probability distribution is not known, hence additionaly we have to adjust optimal portfolio $\boldsymbol{b}$ using maximum drawdown constraint.

$$P(W^{MIN} < \alpha) \leq \beta \tag{11}$$

The probability of our wealth falling below α is at most β. This constraint is approximately satisfied if the following is satisfied, (Busseti et al., 2016):

$$\mathbb{E}[(\boldsymbol{R} \cdot \boldsymbol{b})^{-\lambda}] \leq 1 \ \ where \ \ \lambda = \log(\beta)/\log(\alpha) \tag{12}$$

We hence fit the parameters α, β on a training set and verify on the testing set. The criterion for choosing α, β is maximum final median wealth $median(\boldsymbol{W_F})$ and we only allow 5% of all wealth positions to fall below 90% of the initial wealth.

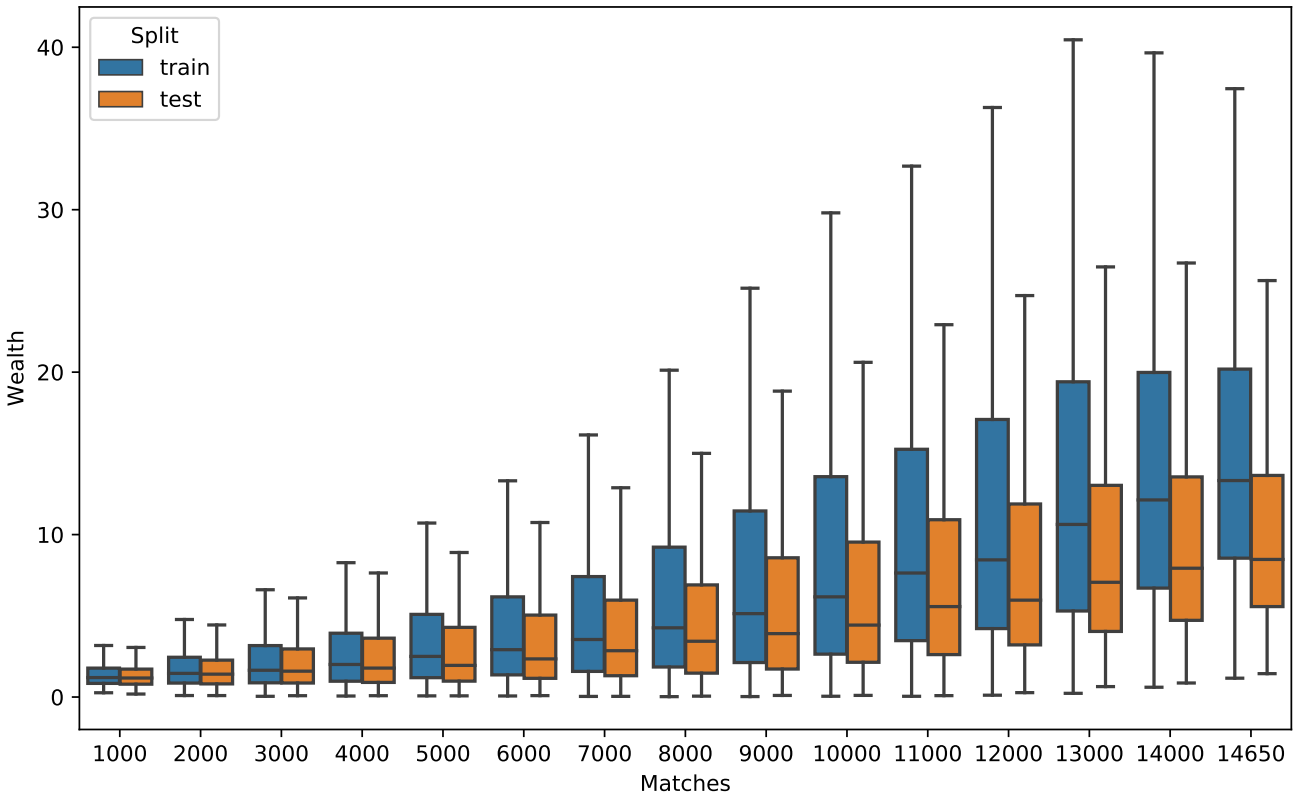


Fig. 2. Results of the risk constrained growth optimal strategy on football dataset.

5. Conclusion

Theoretically the question of the growth optimal strategy has already been answered by numerous great thinkers including (Bernoulli, 2011), (Thorp, 2011) or (Kelly Jr, 2011). It is the Kelly criterion, also known as the geometric mean policy. In practice however, the assumptions of Kelly criterion are rarely met.

In this paper we compare properties of the growth optimal strategy to the traditional modern portfolio theory. Additionally, we showcase it's application in a relevant practical domain with an idea on how to tackle one of it's most difficult challenges: the uncertainty of probability estimates.

References

BERNOULLI, Daniel, 2011. Exposition of a new theory on the measurement of risk. In: *Exposition of a new theory on the measurement of risk. The Kelly Capital Growth Investment Criterion: Theory and Practice*. World Scientific, pp. 11–24.

BUSSETI, Enzo; RYU, Ernest K; BOYD, Stephen, 2016. Risk-constrained Kelly gambling. *arXiv preprint arXiv:1603.06183*.

KELLY JR, John L, 2011. A new interpretation of information rate. In: *A new interpretation of information rate. The Kelly Capital Growth Investment Criterion: Theory and Practice*. World Scientific, pp. 25–34.

MARKOWITZ, Harry, 1952. Portfolio Selection. *The journal of finance*. Vol. 7, no. 1, pp. 77–91.

PETERS, Ole; GELL-MANN, Murray, 2016. Evaluating gambles using dynamics. *Chaos: An Interdisciplinary Journal of Nonlinear Science*. Vol. 26, no. 2, pp. 023103.

THORP, Edward O, 2011. Understanding the Kelly criterion. In: *Understanding the Kelly criterion. The Kelly Capital Growth Investment Criterion: Theory and Practice*. World Scientific, pp. 509–523.

About Authors...

Matej Uhrin was born in Liptovsky Mikulas and is currently studying his first year of PhD at FEL, CTU. The main focus of his studies is machine learning in hybrid domains and his current research is focused on decision making under uncertainty.

Weighting Indicators in Composite Indexes: a Theoretical Model of Competitiveness Assessment for Electronics Market

Mindaugas SAMOŠKA[1]

[1] Dept. of Management, Vilnius Gediminas Technical University, Saulėtekio av. 11, Vilnius, Lithuania

mindaugas.samoska@vgtu.lt

Abstract. *During the last few decades modern companies learned to adapt their strategies to global needs of customers. Leading companies are those that divide and spread their investment volumes, and operating activities more accurate. However, main investment decisions are based on various factors, such as type of the business itself, distance to the main consumer markets, workforce availability, tax rates and etc. These factors in general could be defined as a set of social, economic, geographical and political factors that create a synergy effect while interacting. In general, such interaction's result is known as country's competitiveness level or ease of doing business level. In this case, author is evaluating competitiveness of one particular business sector – companies producing electronics in Lithuania. Article discusses different methods of national competitiveness level assessment, also its' methodological imperfections in weighting competitiveness criteria and possible methods to assess separate business sector's competitiveness, for example, electronics business sector. Correlation-regression and other mathematical programming methods are applied in the research. Result of the research is proposed theoretical quantitative criteria weighting method for national competitiveness assessment models. This model is later adaptable to theoretical model of electronics market competitiveness assessment.*

Keywords

Electronics market, business environment, correlations, indicator weighting.

1. Composite Indexes in Business Environment Assessment

Composite indexes are commonly used in various social studies, also popular science publications to illustrate wider research results in a simply presentable and understandable form. Such indexes are usually fairly generalized and easy to understand, i.e. many different criteria are evaluated and scores are aggregated to simply interpretable one final digital score.

Existing business environment assessment models are based on multi criteria evaluation method usage. Main idea of business growing theory is that any business will eventually decline if it is not expanding constantly. Reason for this is the market competition's pressure.

In order to evaluate correctly basic conditions of possible investment success, companies invest great amounts of money into feasibility studies or market researches. These studies are usually very useful. However, if companies aim is to evaluate basic conditions of business in general in a certain market, they could use and compare results of existing annual competitiveness (ease of doing business) surveys. Numbers of models are used to assess business environment in practice. The two best known and most widely used models are World Competitiveness Report and World Bank's Doing Business Report. They are described more detailed further in the article.

Most annual studies define terms of "business environment", "ease of doing business" and "competitiveness" in a similar manner. Object of evaluation is described similarly as well. Modern economics does not give a unified definition of competitiveness. Terminology in the sphere of competitiveness of the objects under investigation is not standardized [1, 2]. Interestingly, there seems to be no agreed definition of national competitiveness [3]. Since there is no definite margin in terminology, these mentioned terms are used interchangeably in the article as synonyms.

1.1. Competitiveness Research and it's Limitations

It is agreed among scientists that competitiveness research is very specific due to lack of definite and agreed competitiveness term definition [4–14]. However for the same reason scientists have an opportunity to provide their researches with more individual approach towards competitiveness concept and its assessment [14]. Competitiveness potential is often understood as a direct proportion of company's market share compared to company's rivals in private business sector. But it is impossible to use market share as an indicator to assess general country's competitiveness, unless only separate and very specific business sectors (e.g. natural gas exports, electricity market, etc.) were observed [13]. Competitiveness' concept is not the same when talking about private sector companies and all countries economy.

There are two main theories about international competitiveness indicators: first is known as "comparative advantage" and the other is known as "competitive advantage". Country's competitiveness assessment based on international trading indicators is false for another important reason – country's export stats do not include franchising income and foreign venture capital income [10]. High profitability of some business sectors does not necessary mean high country's competitiveness level – Japanese and USA business entities study proved this statement to be true and once more proved that country's competitiveness level cannot be defined nor by terms of separate business sectors profitability nor by country's international trading indicators.

Summarizing World Economic Forum, OECD, European Commission and scientific literature authors [4–14] we state that competitiveness concept described by majority of researched authors and institutions is mostly similar. Most mentioned authors and organizations agree that country's competitiveness is directly linked to country's ability to create liberal and justice conditions to do business. This opinion is also supported by Kao *et al*. [10] – competitiveness is mostly influenced by existing business environment e.g. country's level of ease of doing business conditions.

In conclusion, even though competitiveness, ease of doing business and business environment are often confused as terms used to describe the same phenomenon, most scientists agree that competitiveness level is a result, while ease of doing business is a cause.

1.2. Applying Composite Indexes to Evaluate Competitiveness

Attempts to measure the level of competitiveness began as soon as competitiveness became a widespread acknowledged phenomenon in economics science. At first competitiveness was considered and researched as a simple competitive advantage in micro or macro economy level. Often it was depended on country's accessibility to natural resources. The first attempts to measure the level of competitiveness were based on a mercantilist attitude – based on country's wealth or ability to access that wealth.

In the past century, as the economics science evolved rapidly, competitiveness, business environment or ease of doing business level was started to be evaluated as much more complex phenomenon. Scientists named various competitiveness factors, indicators, also criteria to evaluate these factors. More advanced, accurate data collection methods; advanced methods of data evaluation were introduced.

There are 5 to 10 globally well-known and acknowledged competitiveness assessment methods. Most of them are base for the annual studies and present their results in form of composite index. According to the results of such studies, most of institutes announce index of competitiveness level annually (e.g. World Bank's "Doing Business"; World Economic Forum's "Global Competitiveness Index" and etc.).

1.3. Critical Approach towards Existing Competitiveness Assessment Methodological Potential

All methods of competitiveness assessment obviously have structural similarities. Other technical similarities as well – data sources (official statistics), applied methods (uniform criteria weights; expert evaluation; multi-criteria evaluation; etc.).

Every method has its opponents and its methodological challenges. For instance, World Bank admits itself that "Doing Business" methodologically is limited in scope.

Competitiveness indicators are essential to overall understanding and evaluation of national business environment. World Bank's initiated annual research „Doing Business" evaluates 181 different economies and ranks it in one que according to ease of doing business criteria. This criterion is derived by summing up scores given to 10 different criteria groups. The criteria groups represent different business environment indicators. This

easy-to-use calculation method is based on four liner score aggregation steps.

According to model‘s methodology the main evaluation is being made in very first step – engaging scores by the primary criteria. All primary criteria’s are strictly standardized once again to avoid data misinterpretation. After scores are given according to the primary criteria elementary arithmetical operations are made to aggregate scores to higher level once.

While using the „Doing Business” competitiveness is assessed and the rating of most convenient for business countries is formed by applying the mentioned easy-to-use four step aggregation sequence that is detailed and discussed below:

1 step – each business indicator that is grouped into 10 different business environment groups is given a score. This step of assessing the ease of doing business we name evaluation by the primary criteria. Scores differ among 1 to 181 (according which place the evaluated country‘s indicator is in comparison with all the evaluated countries);

2 step – scores given by every primary individual criteria are summed up to aggregate the score given by each criteria group;

3 step – scores calculated in step to are summed up for every country in the survey and general score for every survey‘s country is aggregated by operations made in the first two steps;

4 step – according to the calculated scores for each survey‘s country (step 3) the general rating is made. Rating is a queue made by the criteria „ease of doing business“ score level.

Local experts are cooperated with and employed in evaluation process. General team of experts is made of more than 3600 local experts that communicate among each other during annual conferences or via electronic means. Such methods are used in order to avoid any possibilities to data misinterpretation and losing survey‘s representation [16].

In calculating the final ease of doing business index score there are no weights given to any of criteria groups. Weighting system in this case is liner weight system [15].

2. Perspectives of Improving Criteria Weighting in National and Sectoral Competitiveness Assessment

Main disagreements among scientists occur in three phases of competitiveness assessment process – object’s conception, indicators and indicator weighting process. And conclusion was drawn that weighting process is the phase that might be improved without achieving common agreement on competitiveness conception (and definition too) neither final list of indicators. Quantitative method of indicator weighting is not developed yet. However based on scientific literature analysis carried by authors, there are theoretical assumptions about such method. A group of scientists from Michigan University – Kumar, Motwani & Stecke [17] – made a list and normalized company’s competitiveness indicators into comparable scale and calculated the proportion of each indicator’s influence to company’s financial indicators. Normalization process was done in assistance of external expert team. In this case scientists achieved a formula that only consisted with names and digits. And the researchers had a final result of indicators interaction – company’s financial results. This led to completely quantitative calculations of each company’s competitiveness indicator which meant that by comparing the starting values and final values of each indicator it was possible to find every indicator’s relative weight by applying quantitative methods only.

Of course, this principal was proposed for solving a one particular problem – calculating company’s competitiveness and finding out how influential to company’s financial results each of stated indicators was. But Kumar, Motwani and Stecke method’s principle is inspiring to develop a quantitative method for ease of doing business indicator weighting process. If we find every indicator’s digital value (score) that was calculated or simply normalized into one unified scale, we could compare every indicator’s score with indicators interaction result if one is possible to calculate. This means that we would have as many formulas (and results) of such divisions as the number of indicators are. Ratio of these results would represent each indicator’s quantitative weight. In such case every individual indicator’s weight would be expressed by a ratio between the normalized individual indicator’s score and digital result of all indicators interaction. Interaction’s score have to be a digital number in order to be sizeable to individual indicator’s score which have to be digital too. The entire indicators interaction scheme is shown in Figure 1.

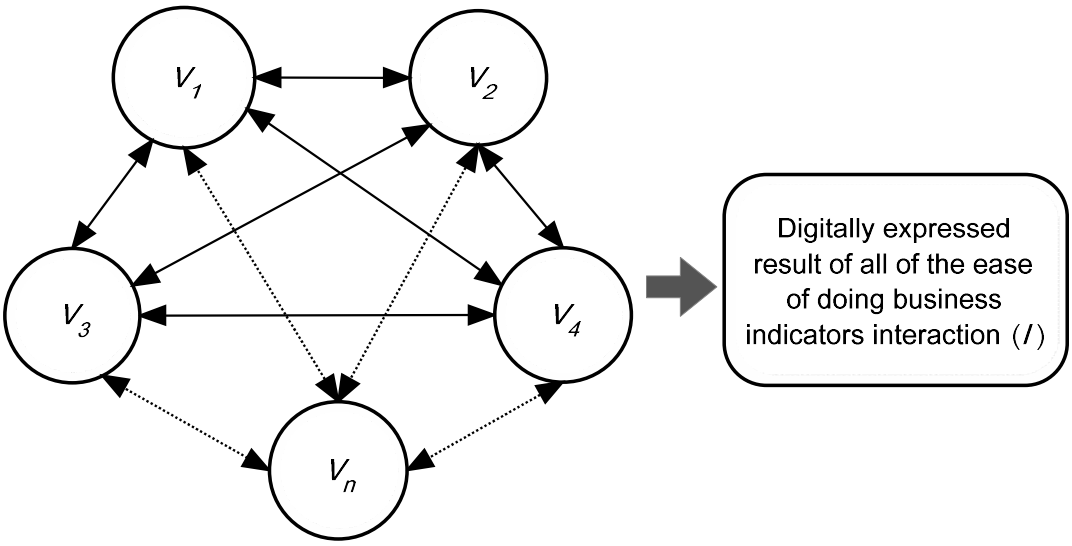


Fig. 1. – Competitiveness Indicators Interaction Scheme. Source: compiled by author.

We suggest an idea that weight of each individual competitiveness indicator could be expressed by this simple formula:

$$w_i = \frac{V_i}{I}, \qquad (1)$$

here: w_i – the weight of ease of doing business i indicator;

I – the result of all of ease of doing business indicators interaction;

V_i – the score of i indicator of ease of doing business.

To calculate competitiveness indicator weights by applying (1) formula it is necessary to have digital data first – normalized scores of each competitiveness indicator and digitally expressed score of all indicators interaction (I). To make formula effective, the unknown (I) have to be calculated by applying quantitative methods only. This condition is also required for competitiveness indicators scores.

3. Theoretical Model for Competitiveness Assessment of Electronics Market

Assessing competitiveness of particular sectors of economy is as important as assessing National Competitiveness. Furthermore, it is also as challenging.

However, different from National Competitiveness, the competitiveness of a chosen particular sector of business, in this case – Electronics market is easier to assess due to relatively smaller number of criteria's. For instance, authors decided to assess competitiveness of Electronics market by using only quantitative criteria. In this case usual composite index formulating tasks occur:

1) Selecting a set of assessment criteria
2) Grouping sets of criteria if necessary
3) Determining methods of scoring criteria
4) Picking method of criteria weighting

These steps of quantitative sectoral business competitiveness assessment are relevant in assessing any business sector's competitiveness. Especially significant part of composite indices is indicator weighting [18-19]. In some cases, indicators weighted based on their relevance to a common result [17, 20-21].

However when assessing Electronics market competitiveness, certain specific details has to be taken into account. Person who is using assessment model must define Electronics market segment, draw its boundaries, so the data were comparable and equally understandable in different counties, leaving a possibility to carry an international research study as well.

3.1. Electronics Market Competitiveness Assessment Stages

Principal shown in Fig.1 inspired the idea of competitiveness assessment model based only on quantitative criteria for Electronics market. The only real problem here is to determine a set of criteria, describing Electronics market that would also meet requirements for criteria described in Fig. 1, i.e. would possibly have a digitally expressed result of all criteria interaction.

In this case, following 4 stages of competitiveness assessment model needs to be implemented: Selecting a set

of criteria, defining Electronics market's competitiveness; determining methods of scoring each criteria – a unified common scale; applying criteria weighting principal (Fig. 1, formula (1)); calculating final score and ranking competitiveness of Electronics market.

3.2. Schematic View of the Model

Since author's developed Theoretical Model for Competitiveness Assessment of Electronics Market is complicated to properly explain on it's own, it is rational to develop a schematic picture. Such visual aid is necessary in understanding stages and steps of employing the model. Models scheme (including indicator weighting stage) shown in Fig. 2.

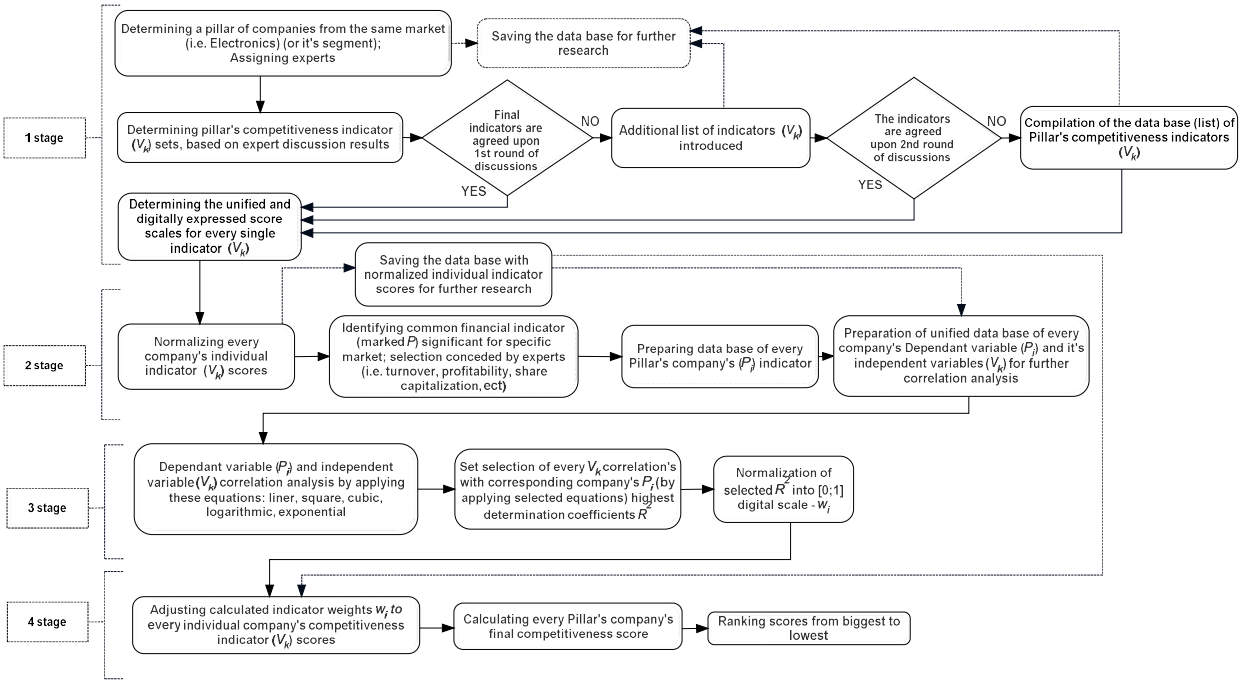


Fig. 2. – The Scheme of Theoretical Model for Competitiveness Assessment of Electronics Market. Source: compiled by author.

Model scheme in Fig. 2 clearly illustrates expert roles in assessment process, a need for data and necessary steps to assess Electronic market competitiveness. Furthermore, author's intention was to develop a model that could use only quantitative date. This way error possibility is minimal.

Advantage of using this model is that the model is self-corrective. It means that even if competitiveness indicators in such dynamic market like Electronics changes rapidly, model is adaptive too. Experts would simply pick a new Pillar of competitiveness indicators that reflect current market situation, and use it in the model.

Furthermore, this model is also adaptable to other business sectors, and based only on quantitative data, so it is as objective and unbiased as it can theoretically be claimed to be.

Acknowledgments

Research described in the paper was developed by author as a part of his research for final paper of his PhD studies.

References

[1] TURKINA, L.; BELOVA, J.; MICKEVIČIENĖ, R. Modeling of ship competitiveness indexes. *Economics & Management*, 2007, p. 1220-1226.

[2] BALKYTĖ, A.; TVARONAVIČIENĖ, M. Perception of competitiveness in the context of sustainable development: facets of "Sustainable competitiveness". *Journal of Business Economics and Management*, 2010, vol. 11 no. 2, p. 341–365.

[3] KRUGMAN, P. Competitiveness: A Dangerous Obsession. *Foreign Affairs*, 1994, vol. 73, no. 2, p. 28–44.

[4] BOSHOFF, C.; TAIT, M. Quality perceptions in the financial services sector: The potential impact of internal marketing. *International Journal of Service Industry Management*, 1996, vol. 7, no. 5, p. 5–31.

[5] LOBANOVA, L. Konkurencingumo modelis: žmoniškųjų išteklių kokybės potencialas. *Organizacijų vadyba: sisteminiai tyrimai*, 2001, vol. 20, p. 135–154.

[6] LEINBACH, T.; BOWEN, J. AirCargo services and the electronics industry in Southeast Asia. *Journal of Economic Geography*, 2004, vol. 4, p. 299–321.

[7] AL-HAWARI, M.; HARTLEY, N.; WARD, T. Measuring Banks' Automated Service Quality: A Confirmatory Factor Analysis Approach, *Marketing Bulletin*, 2004, vol. 16, no. 1, p. 1–19.

[8] KAYIS, B.; KARA, S. The supplier and customer contribution to manufacturing flexibility: Australian manufacturing industry's perspective. *Journal of Manufacturing Technology Management*, 2005, vol. 16, no. 7, p. 733–752.

[9] CLULOW, V.; BARRY, C.; GERSTMAN, J. The resource-based view and value: the customer-based view of the firm. *Journal of European Industrial Training*, 2007, vol. 31, no. 1, p. 19–35.

[10] KAO, C.; WU, W., Y.; HSIEH, W., J.; WANG, T., Y.; LIN, C.; CHEN, L.H. Measuring the national competitiveness of Southeast Asian countries. *European Journal of Operational Research*, 2008, vol. 187, no. 2, p. 613–628.

[11] BENIUŠIENĖ, I.; SVIRSKIENĖ, G. Konkurencingumas: teorinis aspektas. *Ekonomika ir vadyba: aktualijos ir perspektyvos*, 2008, vol. 4, no. 13, p. 32–40.

[12] BALKYTĖ, A.; TVARONAVIČIENĖ, M. Perception of competitiveness in the context of sustainable development: facets of "Sustainable competitiveness". *Journal of Business Economics and Management*, 2010, vol. 11, no. 2, p. 341–365.

[13] MELECKY, L. Approaches to Regional Competitiveness Evaluation in the Visegrad Four Countries. *Mathematical Models and Methods in Modern Science*, 2011, p. 184–189. 978-1-61804-055-8.

[14] NEVIMA, J. Visegrad Four Countries – Case Study of Econometric Panel Data Model for Regional Competitiveness Evaluation. *Journal of Competitiveness*, 2012, vol. 4, no. 4, p. 3–15.

[15] TITARENKO, V.; ZABIELAITĖ, J.; RUDZKIS, R.; ROJAKA, J. Verslo sąlygos ir jų gerinimas Lietuvoje. *Taikomojo mokslinio tyrimo ataskaita*, 2007, vol. 1, p. 132.

[16] MASATLIOGLU, W.; RIGOLINI, J. Labor Dynamics and the Informal Economics: New York, 2005.

[17] KUMAR, A.; MOTWANI, J.; STECKE, K.E. A quantitative approach to measure quality-based competitiveness of an organization. 1999, *Working paper no.* 99-001 (University of Michigan. Business School. Faculty Research).

[18] GINEVIČIUS, R. 2011. A new determining method for the criteria weights in multi criteria evaluation. *International Journal of Information Technology & Decision Making (IJITDM)*, 2011, vol. 10, no. 6, p. 1067 – 1095.

[19] SAMOŠKA, M. Visuomenės gerovės ir verslo sąlygų palankumo vertinimo tyrimų analyze. *Mokslas – Lietuvos ateitis, 2013, vol. 4, no. 1, p. 24-30.*

[20] MELLY,B. Public-private sector wage differentials in Germany: Evidence from quantile regression. *Empirical Economics*, 2005, vol. 30, no. 2, p. 505–520.

[21] VAN DE WALLE, S.; VAN RYZIN, G. The order of questions in a survey on citizen satisfaction with public services: Lessons from a split-ballot experiment. *Public Administration*, 2011, vol. 89, no. 4, p. 1439–1450.

About Authors...

Mindaugas SAMOŠKA was born in Biržai city, Lithuania. For one semester during economics bachelor studies in Vilnius University he had been studying in Matej Bel University in Slovakia. At present, Mindaugas is a third year PhD student at Vilnius Gediminas Technical University in Vilnius, Business Management Department. Mindaugas was a participant of POSTER 2006, POSTER 2009, POSTER 2010 and POSTER 2012 conferences; also Mindaugas is an author of several scientific articles in Lithuania that have been published in reviewed business science journals.

Economic Evaluation of FDM 3D Printing Method

Daniel Běhal[1], *Tomáš Tichý*[2], *Lukáš Dvořáček*[1]

[1] Dept. of Economics, Management and Humanities, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic
[2] Dept. of Electrotechnology, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic
behaldan@fel.cvut.cz, tichyto4@fel.cvut.cz, dvoral14@fel.cvut.cz

Abstract. *The aim of this paper is to make economical calculation of 3D printing technology by FDM method (Fused Deposition Modeling) for electrotechnical industry. The calculation is made for 3 different types of printable materials, filaments (PLA, PET-G, ASA) and is only complementary activity to main business.*

This paper mostly deals with costs of 3D printing itself and does not include costs such as cost of workspace and its maintenance etc. The authors mainly focused on measuring and analyzing costs of power, material, 3D printer maintenance and investment.

Lastly paper compares option of printing cases instead of buying them from electrotechnical market.

Keywords

FDM, 3D printing, PLA filament, PET-G filament, ASA filament, Costs calculation.

1. Introduction

Nowadays there is a great demand for highly specified applications in electronics, which might not be in large production volumes. These applications in the electrical industry often need different types of cases and auxiliary parts. Here comes the FDM 3D printer with low cost and the ability to create complex products from different plastic materials.

The FDM 3D printing method is an additive based method that uses plastic materials in the form of strings of materials (filaments) inserted into the printhead. In the printhead, the filament is melted, and then, by means of a nozzle is applicated to individual layers. For smooth and controlled movement of the print head, a stepper motor system is used to move the printhead in the x, y, z axes [1]. Principle of FDM method is shown on Figure 1.

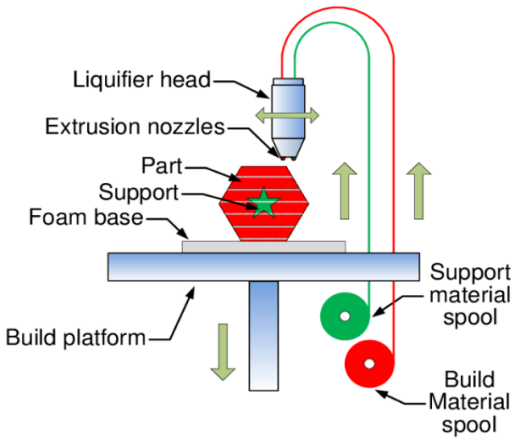


Figure 1. The principle of FDM [1]

2. Problem formulation

As mentioned in the introduction, today's era is known for its incredible speed in innovation. This puts a lot of pressure on the production, cause of frequent changes. Expensive molds for smaller production volumes are huge disadvantage for medium and small businesses, so a suitable solution must be found. The authors saw a possibility of solution in 3D printing, as possible way to rival for medium and small electrical companies.

2.1 Calculation methodology

The authors decided to use model, where 3D printing is a complementary activity, mostly used to aid the production of main product to avoid using subcontracts and to gain more freedom in customization of product. This calculation neglects some costs, such as production building, power fees, but not power cost itself, because all these costs were necessarily already paid to realize production of main product. The calculation was made for three mainly used materials – PLA, PET-G and ASA [2].

It is expected, that the costs of 3D printing will be higher than price of mass-produced mold made products available in a specialized shop, because 3D printing can not beat mold production in terms of time needed to create a product and a price of used materials, that is many times lower then price of used filaments.

3. Economic model

To compare the variants, an economic model was used to calculate the cost of producing a standard box used in electrical engineering. The considered box, which can be purchased in a specialized shop, has parameters 57x90x24mm and its price is 38,97 Kč. The goal of this model is to quantify the cost of manufacturing this FDM box with a variety of materials and compare these costs with the purchase price of the model box.

Investment costs	4 043 Kč	Maintanance costs	2 438 Kč
Operation cost (shift)	100 Kč	Operation cost (year)	24 000 Kč

Table 1. Indirect costs

Printing time (box)	3,00 h	Filament consumption (box)	14,3 m
Printing time (cover)	1,75 h	Filament consumption (cover)	12,4 m
Average yearly production		480 sets	

Table 2. Printing parameters

3.1 Common input costs

Common input costs are cost that have been publicly available, specifically the cost of filaments, 3D printers, and estimated service costs. Investment costs were allocated to equivalent annual amounts using an annuity of 3% discount. The data were obtained from the e-shops focused on 3D printing. Final costs are shown on Table 1 and Table 3.

3.2 Observed input costs

These are data obtained from the created 3D model of already mentioned product from the market. The measurement was performed on a Prusa i3 3D printer MK3S from Prusa Research s.r.o., the company that provided statistical technical data for the calculation. Furthermore, the authors measured the energy consumption when printing various materials, this knowledge is very important for larger plants. Relevant data ae show on Table 2 and Table 3.

	PET-G	PLA	ASA
Filament price	460 Kč/kg	490 Kč/kg	625 Kč/kg
Ratio of defective products	15%	15%	5%
Material costs	31,52 Kč	33,58 Kč	42,83 Kč
Electricity costs	1,76 Kč	2,65 Kč	3,87 Kč
Complete set cost	**111,30 Kč**	**114,70 Kč**	**115,71 Kč**

Table 3. Direct costs

4. Conclusion

From the data collected, it seems that specifically for this application it is better to buy a box from the electrotechnics market. However, if customization were needed for small and medium-sized applications, 3D printing technology would pay off. In addition to this assessment, the costs associated with the operation of the 3D printer were also described.

The following statistical data calculations can be used to read the production of various products with this FDM 3D printing technology.

Acknowledgements

In this section, the authors would like to thank the Department of Economics, Management and Humanities and the Department of Electrotechnology at CTU in Prague for all the support without which the authors would not be able to implement this project. All necessary equipment was funded by a grant No. SGS19 / 062 / OHK3 / 1T / 13.

References

[1] F. Ning, W. Cong, J. Wei, S. Wang, and M. Zhang, 'Additive Manufacturing of CFRP Composites Using Fused Deposition Modeling: Effects of Carbon Fiber Content and Length', p. V001T02A067, Jun. 2015.

[2] Petr Veselý, Eva Horynová, Tomáš Tichý, Ondřej Šefl, 'Study of electrical properties of 3D printed objects', Poster 2018 22th InternAbout Authors...

.About Authors...

Daniel BĚHAL Actually bachelor student on FEE CTU in Prague, in study program Electrical Engineering, Power Engineering and Management, focused on Economics. Theme of bachelor work is on 3D print field

Tomáš TICHÝ PhD student in department of Electrotechnology, FEE CTU in Prague. Actually works on field of electrical properties of 3D printed structures.

Lukáš DVOŘÁČEK PhD student in department of Economics, Management and Humanities, FEE CTU in Prague. Actually works in field of economics in power eingeneering.

Decentralized electricity supplies of remote area

Vladislav Shlapak

Department of Economics, Management and Humanities, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic
shlapvla@fel.cvut.cz

Abstract. *Nowadays, the devices for the operation of which requires electrical energy develop rapidly. In this regard, the question arises of ensuring uninterrupted and high-quality electricity supply to consumers. In Russia, only 30% of remote areas have high-quality energy supply, the rest, and this is almost 23 million people, is provided with electricity by diesel generators. Delivery of fuel to the diesel generator encounters a lot of problems: the lack of railways, a large amount of snowpack, which block roads in winter season. Therefore the average price of 1 kWh is equal 1-3 rub in regular area and 50-100 rub in remote area. One way to reduce the price of electricity in the remote areas is decentralizing of electricity supply.*
One of the types of remote area power supply is a hybrid power plant. To build a hybrid power plant, it is necessary to solve a few of problems:

1. *Determine which types of resources are available and calculate their potential to use for electricity production;*
2. *Design of the scheme of the hybrid power plant;*
3. *Evaluation of efficiency of the hybrid power plant.*

The data obtained as a result of the work can be used to solve a similar problem.

Keywords

Renewable energy sources, photovoltaic power plant, hybrid power plant, remote area, diesel generator.

1. Type of solutions for the power supply in remote areas

A stand-alone power system (SAPS or SPS), also known as remote area power supply (RAPS) - an electrical station, a set of installations, equipment and apparatus used directly for the production of electrical energy, as well as the structures and buildings necessary for this in a certain area, in the case when electrical receivers (consumers) receive power through a radial-trunk transmission line galvanically separated from the rest of the general purpose network. One of the types of remote area power supply is a hybrid power plant. In order to determine the final design of the RAPS, it is necessary to conduct a study of the accessibility of resources from a remote region, as well as to determine which of these resources can provide the necessary indicators for power generation

In this paper we consider the Baykal'skoye village, which is located near lake Baikal. This village has the ability to produce electricity by using of solar radiation and wind.

1.1 Wind power plant possibility of the village

To justify the use of wind energy, it is necessary to calculate the potential of the region:

The approximate value of wind speed at height h can be calculated by the following formula [**1**]:

$$V_h = V_v \cdot (\frac{h}{h_v})^{\alpha} \quad \textbf{(1)}$$

where V_h - wind speed at height h; V_v - wind speed at the height of the vane, standard value 10m; h_v - vane height; α - coefficient depending on the average wind speed at the height of the vane.

For more reliable data of the average value a large set of the statistical data of wind speed in the Baikal village is necessary. Due to the lack of data for more early period, data for 2015-2018 (November) were used. [2]

The average wind speed for each month and the average value for the year:

$$V_{av} = \frac{1}{n} \cdot \sum_{i=1}^{n} V_i \quad \textbf{(2)}$$

where V_i - wind speed at position n; n - position of observation.

Tab. 1. Wind characteristics of investigating object (based on data from [2])

Month	$V_{av},\ m/s$	$V_{20},\ m/s$
January	2,1	2,4
February	2	2,3
March	2,6	3
April	3,1	3
May	3,1	3,6
June	3	3,5
July	2,8	3,2
August	2,8	3,2
September	2,9	3,4
October	2,6	3
November	2,7	3,1
December	2,4	2,7
Total	2,6	3

According calculations the using of wind turbines in the Baykal'skoe village is impractical, since the average value of wind speed at a height of 20 meters less than the recommended 4 m/s. **[1]**

1.2 Solar power plant possibility of the village

Since not every region of Russia can provide power supply for own needs by solar power station, it is necessary to calculate how many of solar insolation there is on the territory of the Baykal'skoe village.

Solar radiation entering an inclined plane taking into account cloudiness is calculated by formula: **[4]**

$$Q_{incl} = S_{incl} + D_{incl} \cdot k \qquad (3)$$

where Q_{incl} – total solar radiation falling on an inclined surface, W/m^2; S_{incl} – direct solar radiation falling on an inclined surface, W/m^2; k – Correction factor which is equal to $(1 - an + 0.38 \cdot n^2)$; a – Coefficient, which depends on the environment (land or sea) and on latitude, which can be taken from **[4, table 5]**; n – The number of clouds in unit fractions (n = 0 for a cloudless sky, n = 1 with absolute cloudiness). Cloudiness is 50% on average in January. [3]

Reflected solar radiation is low in the warm season, because of lack of snow cover, so, it can be neglected [4].

Approximate estimation of declination value of the Sun according to Cooper's formula: [5]

$$\delta = 0.41 \cdot \sin(2 \cdot \pi \cdot \frac{284 + N}{365}) \qquad (4)$$

where, N – the ordinal number of the day in the year counted from the first January.

To determine the height of the sun and the angle of received solar radiation flux on the panel at various angles of inclination of the plane to the horizon according to following formula **[4]**:

$$\cos\theta = \sin\delta \cdot \sin\varphi \cdot \cos s - \sin\delta \cdot \cos\varphi \cdot \sin s \cdot \cos\gamma + \\ + \cos\delta \cdot \cos\varphi \cdot \cos s \cdot \cos\omega + \cos\delta \cdot \sin\varphi \cdot \sin s \cdot \cos\gamma \cdot \cos\omega + \\ + \cos\delta \cdot \sin s \cdot \sin\gamma \cdot \sin\omega \qquad (5)$$

where φ – Latitude of the investigated object, rad; δ – The inclination of the Sun, rad; s – The angle of inclination of the plane to the horizon, rad.

Based on calculations the table of total solar radiation with taking into account a cloudiness was created

Tab. 2. Total solar radiation take into account cloudiness

Month	Total solar radiation take into account cloudiness $Q_c, kWh/m^2$
January	236,328
February	149,28
March	266,568
April	287,112
May	132,648
June	84,36
July	179,016
August	84,768
September	142,992
October	285,384
November	280,584
December	92,088

1.3 Load Graphs

According to typical load graphs for each season for the rural area, I calculated loads graphs for Baykal'skoe village

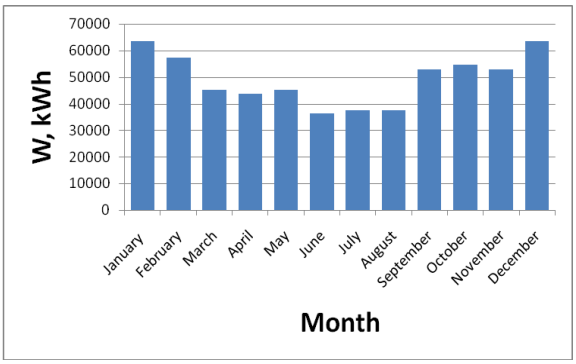


Fig. 1. Total annual active power load graph

2. Design of the scheme of the hybrid power plant

Taking into account load graphs and previous calculation of solar radiation I chose inverters, number of solar panels, type and number of batteries, solar controller.

The equipment's parameters were tabulated.

Tab. 3. The main elements of a hybrid power plant

Name of equipment	Cost for 1 unit	Number of equipment	Total cost
ECLIPSE SRP-320-E01B	18,500	504	9,324,000
OPzV 3000 (12V / 200Ah)	113,477	63*24	171,577,224
PROSOLAR SUNSTAR MPPT SS-40CX	52,700	7	368,900
MAP SIN "Energiya"	162,800	7	1,139,600

3. Evaluation of efficiency of the hybrid power plant.

One of the ways to increase the efficiency of a hybrid power plant is the selection of the optimal number of batteries and solar panels. The basis of the selection of numbers of solar cells and accumulator batteries is the problem of linear programming, where the number of solar panels and number of accumulator batteries, and accordingly costs, is minimal. By solving this problem is possible to find the minimal and cost-effective amount of panels and batteries.

Using the data were obtained in previous calculations, we can determine the objective function. The price of solar panels and the price of batteries should be minimized.

$$C_{pan} \cdot N_{pan} + C_{bat} \cdot N_{bat} \Rightarrow MIN$$
$$18500 \cdot N_{pan} + 2723448 \cdot N_{bat} \Rightarrow MIN \quad (6)$$

Where C_{pan} - cost of panel; N_{pan} - the number of panels; C_{bat} - cost of battery; N_{bat} - the number of batteries.

It is also necessary to take into account that the batteries should not be discharged more than 70% for extending their durability. And also batteries should not be charged more than 100%. It will be two main constraints in the minimization problem.

For an example of calculation I took April 2018. Firstly, after dividing the day into temporary parts of 3 hours each, the required amount of constraints were found:

$$N_{row} = \frac{24}{N_{hour}} \cdot N_{day} = \frac{24}{3} \cdot 31 = 248 \quad (7)$$

where N_{day} - amount of day in a month; N_{hour} - the size of the period of separation of the day

For each parts of the day, I must use two constraints. That means I would have approximately 500 constraints as a total.

The formula to determine the battery charge at time t + 1:

$$W_{batt(i)} = W_{gen(i)} + W_{batt(i-1)} - W_{cons(i)} \quad (8)$$

where $W_{batt(i)}$ - charge of battery at time t, $W_{gen(i)}$ - amount of generated energy by solar panels, $W_{batt(i-1)}$ - charge of battery at time t-1, $W_{cons(i)}$ - amount of consumption energy at time t.

The constraints have the form:

$$\sum_{n=1}^{248} P_{gen} \cdot N_{pan} \cdot t_{(n)} - \sum_{n=1}^{248} W_{cons(n)} \le 0$$
$$P_{gen} \cdot N_{pan} \cdot \sum_{n=1}^{248} t_{(n)} \le \sum_{n=1}^{248} W_{cons(n)} \quad (9)$$

where P_{gen} - generated power by one solar panel; N_{pan} - amount of solar panels; W_{cons} - amount consumption energy.

$$\sum_{n=1}^{248} P_{gen} \cdot N_{pan} \cdot t_{(n)} - \sum_{n=1}^{248} W_{cons(n)} + 0,7 W_{batt(0)} \ge 0$$
$$P_{gen} \cdot N_{pan} \cdot \sum_{n=1}^{248} t_{(n)} + 0,7 \cdot E_{nom.bat} \cdot N_{bat} \cdot U_{bat} \ge \sum_{n=1}^{248} W_{cons(n)} \quad (10)$$

where W_{batt} - charge of the battery; $E_{nom.bat}$ - capacity of the battery; U_{bat} - a voltage level of the battery.

The solution of this problem is the number of solar panels and the number of batteries at the lowest cost and ensuring of the consumers' needs.

The number of solar panels is 520 units.

The number of batteries is 50 units

Tab. 4. Results of optimization

	Without optimization	With optimization
N_{pan}	504	520
N_{bat}	63	50
Cost, rub	180,901,224	145,792,400

In conclusion, we can say that using optimization methods we managed to improve the result by 35 108 824 rubles (19,40%).

4. Results of research

1. In remote areas such renewable source of energy as solar radiation and wind energy have good prospects for development
2. Based on the calculations, only solar radiation meets the standards for the production of electricity.
3. Taking into account the fact that in some months electricity production exceeds consumption, it was decided to install batteries.
4. One of the ways to improve the efficiency of a hybrid power plant is to change the angle of rotation of the solar panels relative to the sun or to choose an effective number of solar panels and batteries.
5. To choose an effective number of solar panels I used a linear programming task. The solution to the minimization problem is the number of solar panels in the amount of 520 pieces and the number of batteries in the amount of 50 pieces. This will cover the required amount of electrical energy, as well as reducing the required investment by 19.4%.

Acknowledgements

Research described in the paper was supervised by Prof. T. Králík, FEL CTU in Prague

References

[1] LUKUTIN B. V. *Vozobnovlyayemyye Istochniki Energii: Uchebnoye Posobiye – Renewable Energy Sources: Study Guide*. Tomsk: Tomsk Polytechnic University, 2008.

[2] GISMETEO. Dnevnik pogody v Baykal'skom – The Weather Diary of Baykal'skoe Village. Available at: https://www.gismeteo.ru/diary/198537/2015/2/ [Accessed: November 20, 2018]

[3] *Dnevnik pogody v Baykal'skom – The Weather Diary of Baykal'skoe Village*. Available at: https://rp5.ru/ [Accessed: January 15, 2019]

[4] SAMOYLOV D. V. *Raschet velichiny Postupleniya Teploty ot Solnechnoy Radiatsii na Poverkhnost' Zemli: Metodicheskiye Ukazaniya – Calculation of Incoming Heat value on the Earth Surface from Solar Radiation: Methodical Guide.* Moscow: Bauman Moscow State Technical University, 2006.

[5] Matthias Gunther "*Advanced CSP Teaching Materials". Chapter 2 "Solar Radiation*". [Online]. Available: http://www.energyscience.org/bibliotheque/cours/1361469594Chapter%2002%20radiation.pdf [Accessed: January 15, 2019]

About Authors...

Vladislav SHLAPAK was born in Norilsk city, the Russian Federation. He received a secondary education in physical and mathematical school, higher education in the Tomsk Polytechnic University. He is a bachelor in electric power industry.

Development of a hybrid power plant for the power supply of a rural area

Alena PAVLOVA

Dept. of Economics, Management and Humanities, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

pavloal2@fel.cvut, aluonka95@yandex.ru

Abstract. *The aim of this work is to develop a decentralized hybrid power plant for the rural area. As an example, I chose the village of Nazino. Since this place is located away from major cities and its electricity consumption is rather small (about 400 people live in the village, there are no large enterprises) it is impractical to build a power line and connect to the general power grid. Therefore, at the moment the village receives energy from the diesel substation. However, this option of electricity generation cannot be called optimal. The problem is that the village is located in a hard to reach place, and taking into account all transportation costs for the delivery of diesel fuel, the cost of electricity becomes quite expensive. Using an alternative energy source will not only help improve the environmental situation at the substation, but also lead to significant fuel savings. In the case of building a hybrid substation, the project will initially require significant capital expenditures, but then the variable costs will be extremely small. In this paper, I developed several scenarios with various substation components, evaluated the selected scenarios from an economic point of view, and also conducted a sensitivity analysis.*

Keywords

Hybrid power plant, renewable energy, solar energy, financial assessment, sensitivity analysis

1. Introduction

In our time, the power supply system in Russia is well developed and meets the standards of electricity quality. The average cost of electricity is 1,13 czk/ kW [1]. In the ranking of European countries with the cheapest cost of electricity Russia is on the 3rd place [2]. However, there are a number of features that overshadow the energy situation in Russia.

One of these feature associated with the fact that Russia occupies a huge territory - 17 125 191 km^2 [3], but the population is not evenly distributed. And, accordingly, there are many settlements with a small population that are quite remote and distant from large cities.

For such places, it is economically inexpedient to build overhead power lines and connect them to the General power system, because it requires huge capital investments. More acceptable, and the most common today is the option of using diesel power plants. However, in this case, the cost of electricity will be quite high (can reach 36,8 czk/kW). This is due to the high cost of diesel fuel and its delivery to the place of use.

In this paper work, I will consider the case when the connection to the centralized power supply systems is not economically justified, and the delivery of fuel makes the cost of electricity too high. In this situation, it is very profitable to use a hybrid power supply system, in which electricity received not only from traditional sources, but also from alternative ones.

2. Sun as alternative energy source

Now the use of solar panels is quite common, which not only speaks about the effectiveness of their use, but also carries additional advantages, such as: the abundance of the necessary information about the panels, wide range of equipment on the market, a large number of specialists in this field.

Also, in addition to this, solar panels have the following advantages:

1) Low Maintenance (Solar panels have no moving parts and require very little maintenance beyond regular cleaning. Without moving parts to break and replace, after the initial costs of installing the panels, maintenance and repair costs are very reasonable.)

2) Environmental friendliness (The sun is both a renewable and clean source of energy. In addition to the fact that the sun energy does not create greenhouse gases, such as oil based energy does, nor does it create waste that must be stored, such as nuclear energy, it also helps to reduce global warming. The sun is really an endless source of energy. And is present in varying degrees at every point on the globe. For example, as we can see on the map taken from the source [4], in the village of Nazino the value of solar insolation reaches 3.5-4 kWh/m^2day.)

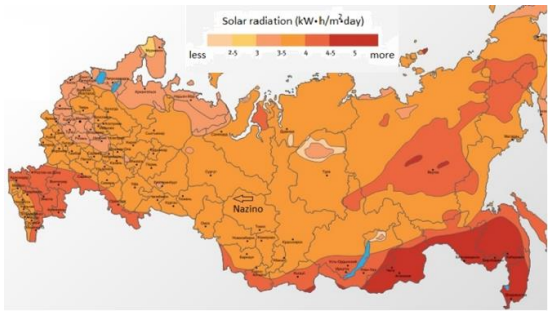


Fig.1. The value of solar radiation in Russia [8]

3. Technical solution

As everyone knows solar panels produce energy only in the hours of solar activity, they cannot fully cover the necessary load (there will be a lack of energy at night or in cloudy hours). Therefore, solar panels should be used either with batteries or with another energy source, such as diesel generators.

In the case of joint operation of batteries and solar panels, in the hours of solar activity, the batteries will not only cover the necessary load, but also charge the batteries, which in turn will give out energy in those hours when the energy from the panels is not enough or it will be completely absent.

Almost the same logic will be applied to the system with diesel generators. In those hours, when the energy from solar panels will be enough or it will be completely absent, the necessary load will be covered by diesel generators.

So we have several different variations on how we can design a substation. In this chapter I want to analyze several different scenarios.

All scenarios will be developed with the following conditions:

1. All scenarios will use the same type of solar module - FSM 320P tilted at an angle of 45 degrees.
2. In all scenarios, one type of battery will be used- PROSOLAR OPZV – 3000.
3. Also for all scenarios will be used the same type and power of diesel generators (AD80-T400).

I calculated the amount of electricity produced by the selected module using the site [5]. This information shown in figure 2.

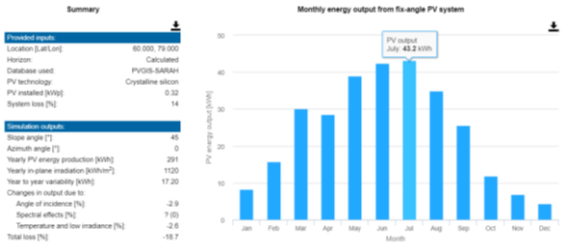


Fig.2. Monthly power generation by 1 module FSM 320P[5]

Now I want describe the scenarios.

1.) Because the Sun is the most active in July (as shown in Fig. 2), we design the substation so that this month all the necessary energy will be generated by solar panels, which will work together with the batteries. In other months, solar insolation is lower, so the energy produced by the panels will not be enough to supply electricity. At such moments we will additionally connect a diesel generator. In the morning, during peak electrical loads, we can turn on diesel generators with solar panels, cover the necessary load and charge the batteries. Then, when the peak consumption will pass, we will use the stored energy, and if necessary, turn on the generators in the evening peak loads.

2.) Since batteries are very expensive compared to other equipment we will try to minimize their number. We will design a substation that will consist of solar panels, diesel generators and batteries. But unlike the first option, the generators also will be connected in July to minimize the number of batteries. We will use generators as in the first scenario, covering the necessary lack of energy during peak hours of loads and charging the batteries.

In this scenario, the number and type of PV panels and generators will be the same as in the first scenario, and we will calculate the necessary amount of batteries according to constraints (rechargeable batteries are always charged at a level from 30 to 100%).

3.) In this scenario, the substation will consist entirely of solar panels and batteries. Thus, since solar insolation is minimal in December and maximum in July, the amount of equipment will be chosen so that the panels could in the summer, to store such an amount of energy, which would be enough for the winter months with a smaller level of insolation.

4.) To compare the designed substation with the traditional substation I will also consider in my work the option without connecting panels. In this version (as in the real substation in the village of Nazino) all loads are covered exclusively by diesel generators and batteries. The use of batteries will help to use the generators most energy-efficient.

Assuming that consumption is maximum in December, I will select batteries so that I can cover the entire load without generators (only with the help of batteries).

5.) In this scenario, the substation operates exclusively at the expense of diesel generators. Electricity is covered by diesel generators as needed.

6.) In this scenario, the substation will run on solar panels and with diesel generators, but unlike the first two scenarios, we will not use batteries.

For better clarity I created the table which show us which equipment we will use in all scenarios.

	Solar panels	Batteries	Diesel generator
1 scenario	+	+	All year, except July
2 scenario	+	+	+
3 scenario	+	+	-
4 scenario	-	+	+
5 scenario	-	-	+
6 scenario	+	-	+

Tab.1. Description of equipment which we will use

For each scenario, I made calculations on the choice of the amount of necessary equipment. And also in the variants where diesel generators are used, I calculated the consumption of diesel fuel (for further economic calculations). The results of the work you can see in table 2.

Scenario	Number of solar panels	Wpv [kWh]	Number of batteries	Wgen [kWh]	Gg(Total) [kg]
1	1 005	292 213	10	335 331	88 084
2	1 005	292 213	4	336 020	88 119
3	2 175	632 401	1958	–	–
4	–	–	4	628 431	162 610
5	–	–	0	630 516	171 557
6	405	117 757	0	512 806	147 087

Tab.2. The result of the calculation for all scenarios

4. Financial model

Now that I have done all the technical calculations and received all the necessary data, I can go to the economic part of my work.

In my economic part of the work, I will consider the following tasks:

1. The point of view of the owner of the substation. In this case, I will compare the scenarios with the fourth one, which reflects the existing situation at the substation.

2. Point of view of the investor. I will consider the case when I install solar panels and sell the generated energy to the substation. At the substation, part of the energy is produced by diesel generators, and some are bought from the me.

4.1. Data for the financial model

4.1.1 Solar panels

The solar cell manufacturer [6] states that:

• preservation of the declared capacity of more than 90% of the rated capacity is guaranteed for 10 years

• preservation of the declared power of more than 80% of the minimum rated power - for 25 years

Thus, to take into account the possible decrease in the performance of solar panels, we will increase the number of panels by 10% for 11 years of use. In my model, this will be reflected in the form of additional investments at the end of 10 year of operation.

For clarity, I will collect all the data on the solar panels that will be used in the model in Table 3.

Scenario	Number of solar panels	Number of additional solar panels	Investments for the zero year [rub]	Investments for the tenth year [rub]	Lifetime [years]
1	909	91	11 535 210	1 154 790	25
2	909	91	11 535 210	1 154 790	
3	1 658	166	21 040 020	2 106 540	
4	0	0	0	0	
5	0	0	0	0	
6	909	91	11 535 210	1 154 790	

Tab.3. Initial data for solar panels

4.1.2 Storage battery

Since I will compare the project that has already been put into operation with new ones, and 542 batteries have already been installed at the power plant. In the model, I will consider the number of batteries equal to:

$$N_{batt}^{\text{mod}} = N_{batt}^{real} - 4$$

where N_{batt}^{real} - the number of batteries required for the substation;

4 – the number of batteries already installed in the substation.

Scenario	Number of batteries	Number of batteries in the model	Investments for the zero year [rub]	Lifetime [years]
1	10	6	140 645 760	15
2	4	0	0	
3	1958	1954	5 447 954 880	
4	4	0	2 242 058 880	
5	0	0	0	
6	0	0	0	

Tab.4. Initial data for storage battery

4.1.3 Diesel generators

Since I will compare the already commissioned project with the new ones, the price of diesel generators will not be included in the model, because the substation

already has them. And in all the scenarios in which they provide for the type of generators is the same as in the existing substation.

The price of diesel fuel in 2017 amounted to:

$$C_{fuel} = 44\ 535\ rub/ton = 44{,}5\ rub/kg$$

However, it is necessary to take into account that the village is located in a remote place and far from major cities. Therefore, we need to find the price of fuel, taking into account transportation costs. Since the payback period has already passed for 2017, we can assume that the cost of electricity consists solely of the cost of diesel fuel consumption and repair costs. I will accept the assumption, and I will assume that 90% of the cost is spent on the purchase of diesel fuel.

According to information from the technical data of diesel power plants of the Tomsk region [7] the approved electricity tariff is T = 21.53 RUB / kWhour. Previously, the value of the total annual electricity consumption (776 MWh) was found. Thus, for the year the production of electricity is spent:

$$C_{year} = W_{year} \cdot T = 628\ 560 \cdot 21{,}53 = 13\ 532\ 897\ rub$$

As noted earlier, the fourth scenario reflects the situation in substation in 2017. Thus, we know that in 2017, at the substation in the village of Nazino, the following was consumed 162 610 kg of fuel. Thus, the cost of diesel fuel including delivery is:

$$C_{fuel.tot} = \frac{0{,}9 \cdot W_{year} \cdot T}{G_{year}} = \frac{0{,}9 \cdot 628\ 560 \cdot 21{,}53}{162\ 610} = 74{,}9\ rub/kg$$

4.2 Economic parameters

4.2.1 Tax

Since this program is the state, therefore, the economic model will not contain taxes.

Based on this fact, the financial model will also not contain the depreciation, as it is used to reduce the amount of taxes paid.

4.2.2 Inflation

Since the project is calculated for 2017, we will find inflation as the average of the inflation values for 2017 and 2018 [8] and the projected inflation for 2019 - 2023 [9]. The calculations are presented in the table below.

Years	Inflation [%]	Average value [%]
2017	2,5	3,67
2018	4,3	
2019	4,2	
2020	4	
2021	3,8	
2022	3,4	
2023	3,5	

Tab.5. Inflation

4.2.3 Discount rate

Since the substation being developed is a state project. I will find a discount rate according to the Russian Federation Decree №1470 [10] according to the formula:

$$d = d_i + \frac{P}{100}$$

where d_i - risk-free discounting factor

$\frac{P}{100}$ - risk adjustment.

The discount rate without taking into account the risk of the project is determined by the following formula [10]:

$$d_i = \frac{1 + \frac{r}{100}}{1 + \frac{i}{100}} - 1$$

were r - refinancing rate,

i- inflation rate.

The key rate was introduced by the Bank of Russia on September 13, 2013 as the main indicator of monetary policy. And for 2017, it is 8.8%. Thus, the risk-free discount rate will be:

$$d_i = \frac{1 + \frac{8{,}8}{100}}{1 + \frac{3{,}67}{100}} - 1 = 4{,}9\ \%$$

Risk adjustment for investment in the intensification of production is 5 percent [10].

Thus, the discount rate will be:

$$d = 4{,}9 + 5 = 9{,}9\ \%$$

4.3 Results of the calculated model, owner point of view

When the substation is reconstructed, the owner will not change the cost of electricity, so the profit from the sale of electricity will remain unchanged. The profit from the reconstructing of the substation will be that it helps the owner save on variable cost (on fuel).

Below in the table I calculated how much money the owner will save annually on fuel under different scenarios of building a substation. From table 30 we know that without making changes to the work of the substation, it annually spends on the purchase of fuel:

$$C_{fuel(4)} = 12{,}18\ mln.rub$$

Subtracting this number from the cost of fuel for each scenario, we get annual savings.

Scenario	Gg(Total) [kg]	Fuel costs [rub]	Annual savings [rub]
1	88 084	6597 492	5 581 997
2	88 119	6 600 113	5 579 376
3	–	–	12 179 489
4	162 610	12 179 489	0
5	171 557	12 849 619	-670 130
6	147 087	11 016 816	1 162 673

Tab.6. Scenario calculation results

As can be seen from the table above, in 5 scenarios the fuel is not saved, therefore, this project will contain only expenses (for the purchase of new equipment). Therefore, at this moment we can already consider that this project is not profitable and will not consider it. If we calculated the economic model built on the 5 scenario, the NPV would be negative.

Having calculated the 1,2,3 and 6 scenarios according to the methodology described earlier, I obtained the results presented in table 7.

Scenario	Number of solar panels	Number of batteries	NPV, rub	Payback period
1	1005	10	17 217 576	8
2	1005	4	34 617 036	3
3	2175	1958	-5 598 424 414	-
6	405	0	3 435 253	8

Tab.7. Results of the calculated model

From the results it is clear that the second scenario is the most cost-effective, then comes the first one, then the sixth one. The third scenario is absolutely not profitable.

4.4 Results of the calculated model, investor point of view

In this model, I will consider the project from the point of view of the investor. The village of Nazino is already provided with energy in the required amount. My project will consist in the fact that we put the solar panels and will sell energy to the substation. The buyer agrees to the transaction, provided that it will be profitable for him (the price for which they will buy energy from us must be lower than that they spend on diesel fuel).

Since this project will not be the state, as the previous one, we need to recalculate the discount rate, take into account taxes and depreciation.

The tax rate is set in article 284 of the Tax Code of the RF. Tax rate is equal to 20%. This value has not changed since 2009. Therefore, in this model it will be assumed that the next 15 years the tax rate will not change and will be 20%. To reduce the taxes paid, I will use linear depreciation.

For finding the discount rate for a new capital investment I will use capital asset pricing model. I will find a discount rate by the formula:

$$r = r_f + \beta \cdot (r_m - r_f)$$

where r_f – risk free rate

r_m – market return

β – beta, which measures the sensitivity to market movements.

From a site containing various useful data for economic evaluations I received the following information:

Equity Risk Premium for Russia:

$$r_m - r_f = 9,43\%$$

Average Levered Beta for Green & Renewable Energy:

$$\beta = 0,88\%$$

According to the Central Bank of Russia, the yield of bonds with a maturity of 20 years is 8,66 %. I will take this percentage as risk free rate:

$$r_f = 8,66\%$$

Substituting all the found values in formula above, I found a discount rate:

$$r = 8,66 + 0,88 \cdot (9,43) = 16,96\%$$

In order to find the cost of electricity for the investor, I need to find the cost of electricity for the owner of the substation. The owner will agree to buy energy from the investor, if it is less than the cost that he spends on the production of electricity.

Thus, the price for electricity will be:

$$C = \frac{C_{fuel.total} \cdot G_{gen}}{W_{gen.year}} = \frac{74,9 \cdot 162\ 610}{628\ 560} = 20\ rub/kWh$$

Thus, from the point of view of the investor, the price for which it will be possible to sell energy should be less than 20 rubles.

Take the cost of electricity equal to 20 rub/ kWh. We know that solar panels produce 292 213 kWh per year. So we will find the revenues for the first year of using solar panels.

$$REV = 292\ 213 \cdot 20 = 5\ 844\ 260\ rub$$

After that I found depreciation. Using the investment information, I received a depreciation value equal to:

$$DEP = 17\ 126\ 950 \cdot \frac{1}{15} = 1\ 141\ 797\ rub$$

I used depreciation to reduce taxes paid, because the value of which I pay tax is the difference from my revenues and my expenses (depreciation), so in the first year the amount of taxes is:

$$Tax = r_{tax}(REV - DEP) = 0,2(5,8 - 1,1) = 0,94\ mln.rub \quad (84)$$

where $r_{tax} = 20\ \%$ – Tax rate.

Then I found the cash flow, deducting the amount of taxes paid from the revenues. For example, for the first year, the cache flow value is:

$$CF = REV - TAX = 5{,}8 - 9{,}4 = 4{,}9 \ mln.rub \quad (97)$$

Then, I got the NPV value:

$$NPV = 11\,116\,756 \ rub \quad (98)$$

After creating the model, I found the minimum price for the investor. For this, I used the "search for a solution" function built into Excel and found the price at which the NPV would be zero.

The minimum cost of electricity is:

$$C = 12 \ rub/kWh \quad (99)$$

5. Sensitivity analysis

Below is the dependence of the NPV project on the price of electricity. As can be seen from the figure, the higher the price, the higher the NPV, however, it should be remembered that the maximum price that can be set for electricity is 20 rubles/ kilowatt hour.

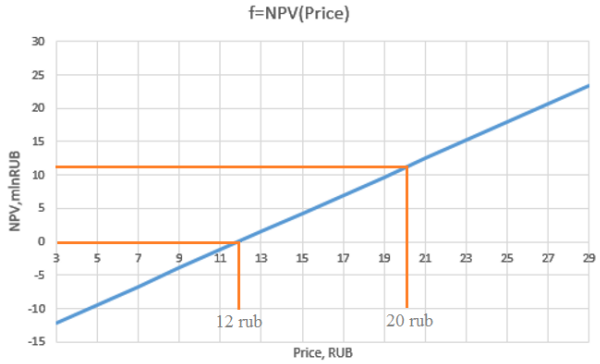


Fig.3. Sensitivity analysis for NPV (Price)

Based on the schedule of dependencies on the price, we can conclude that the project will be profitable if we sell electricity for a price that exceeds 12 rubles per kilowatt hour.

Conclusion

It is indisputable that the use of alternative energy sources has a beneficial effect on the environment. Within the framework of the situation studied in this project, the use of solar panels also had a beneficial effect on the economic component.

Comparing the various options for the substation, the following dependencies were identified:

1. The project is more profitable, when lower amount of batteries is used. This is due to the fact that the price of batteries is much higher than the cost of other equipment (100 times more than the cost of solar panels).

2. Other things equal, the use of solar panels makes the project more profitable.

Thus, if manufacturer applies the project of restructuring the existing substation and rebuilding it into a hybrid power plant (2 scenario), this will increase the project's NPV by 35 million rubles.

References

1. Ranking of European countries on the availability of electricity in 2018. [Online] Available: http://riarating.ru/countries/20180625/630098114.html

2. Ranking of countries on electricity prices-RIA news 19.11.2015. [Online] Available: https://ria.ru/infografika/20151119/1321996028.html

3. Russia - Wikipedia. [Online] Available: https://en.wikipedia.org/wiki/Russia

4. Coal against solar panels. [Online] Available: https://m.aftershock.news/?q=node/296371

5. Photovoltaic Geographical Information System (PVGIS) - European Commission. [Online] Available:http://re.jrc.ec.europa.eu/pvg_tools/en/tools.html

6. Solar modules sunways . [Online] Available: https://s-ways.ru/products/solnechnye-moduli-sunways-serii-fsm/solnechnyy-modul-sunways-fsm-320p.html

7. Tariffs for electricity for Tomsk and Tomsk region . [Online] Available:https://energo-24.ru/tariffs/electro/2018-elektro/12873.html

8. Inflation rate in Russia. [Online] Available:https://bankirsha.com/UROVEN-INFLYACII-V-ROSSIYSKOY-FEDERACII-PO-GODAM.HTML

9. Forecast of inflation in Russia . [Online] Available:https://apecon.ru/prognoz-inflyatsii-na-gody-v-rossii

10. Postanovleniye Pravitelstva Rossiyskoy Federatsii №1470 Ob utverzhdenii poryadka predostavleniay gosudarstvennih garantiy na konkursnoy osnove za schet sredstv Budgeta razvitiay Rossiyskoy Federatsii

About Authors

Alena Pavlova was born in 1995 in Russia. She graduated from Tomsk Polytechnic University with a Bachelor's degree "Electrical Power Engineering". In 2017 she started his Master's degree in "Electric Power Generation and Transportation" at Tomsk Polytechnic University. Through a joint double-degree program she is doing her Master's degree in "Economics and Management of Electrical Engineering" at Czech Technical University in Prague

Hybrid Power Supply in the Koyda Village

Andrey KURGANOVSKIY

Department of Economics, Management and Humanities, Czech Technical University in Prague,
Technická 2, 166 27 Praha, Czech Republic

kurgaand@fel.cvut.cz

Abstract. *The aim of this paper is to integrate renewable energy sources in villages on the arctic side. Most of locations on the arctic side in Russia are not connected to the central network and have expensive power supplied by diesel generators. These locations, for example, the Koyda village, have rich wind energy potential. So, integration of wind turbines with diesel power system may decrease energy cost. In this master thesis, I will research renewable energy potential in the Koyda village, consider integration of wind turbines with diesel power plant and make techno-economic analysis. I suggest obtaining more cost efficient power plant.*

Keywords

Renewable energy sources, wind energy, hybrid power system, financial evaluation.

1. Introduction

Nowadays the environmental problems such as global warming and air pollution have a global scale. The stock of fossil fuels like gas, coal and oil is diminishing every year. These factors touch all counties in the world and because of that the "green" power engineering is developing. "Green" power engineering uses renewable energy sources: sun, wind, water flow, geothermal source and energy of bio-fuel.

Full or fractional changeover to the renewable energy sources has a range of economic and ecological advantages. According to Lukutin et al. (2015) this range includes:

- Speed-down the growth of negative anthropomorphic impact on the environment and preventing climate change taking in account necessity to satisfy the growing energy needs.
- Reducing of air pollution in the way of decrease the fossil fuel burning. Therefore, there will be heath maintenance of the population, decrease breathing system diseases risk, decrease state spending on health maintenance.
- Increase the supply decentralization, therefore its safety and reliability will increase.
- Decrease costs on transmission and distribution the power and fuel and also losses in these processes.[1]

In this paper, I analyze geography of the Arkhangelsk Region and the Koyda village where the power plant is proposed to be build. The geographical analysis includes:

- Evaluation of climate conditions
- Evaluation of solar energy potential
- Determining of wind power characteristics on the territory

After studying natural conditions I analyze consumers' characteristics. Based on these characteristics I will choose the corresponding equipment for power plant. Next, I consider the hybrid power plant and analyze its economic effectiveness.

2. Renewable energy sources in the region

Nowadays, the part of renewable energy generation in Russia is very small but it has a great potential. The outlook of cost-efficient renewable energy usage in Russia consists more than 30% of annual power consuming. However current part of renewable energy is only 0.06%. To compare, Germany generated 151.57 billion kWh with renewable power plants in 2017. It is 37.5% of all generated power. [2]

Description of the Koyda village

The Koyda village is located on the north of Arkhangelsk region in the Mezenskiy district in Russia. Its population is 478 people, according to data of 2017 year. Nowadays, the village receives electric energy from diesel generators only. Its total capacity is 360 kW. It consumes 360 tons of fuel per year. The fuel is shipped by sea [3]. Integration of renewable energy sources in the power supply system will help to decrease costs on fuel and will decrease CO_2 emission. There are no rivers with dramatic water flow in the Koyda village, so, next, I will analyze solar and wind energy potential in the region.

Potential of solar energy in Koyda

I take the data about Archangelsk because it is the nearest point to the Koyda village where the data exists. According to "www.solbat.su" the year insolation of Arkhangelsk is 0.85 MW. The monthly insolation data are in Table 1.

Table 1 - Monthly insolation in Arkhangelsk [4]

Month	Jan	Feb	Mar	Apr	May	Jun
Insolation (MJ/m^2)	12	61	207	356	494	575
Insolation (kWh/m^2)	3.3	16.9	57.5	98.9	137.2	159.7
Month	Jul	Aug	Sep	Oct	Nov	Dec
Insolation (MJ/m^2)	565	385	186	71	20	4
Insolation (kWh/m^2)	156.9	106.9	51.7	19.7	5.6	1.1

According to Lukutin et al.(2015), the photovoltaic power plant will be effective if the yearly insolation per square meter is more than 1000 kW*h. If we calculate the annual insolation in the Koyda village presented in Table 1 than we will obtain 815.4 $kW*h/m^2$. This number is less than 1000 kW*h, therefore it is not efficient to install solar power plant in the Koyda Village.

Potential of wind energy in Koyda

There is monthly average wind velocity in Table 2. The data are from "www.rp5.ru".

Table 2 – Monthly wind velocity in Koyda [5]

Month	Jan	Feb	Mar	Apr	May	Jun
Wind velocity(m/s)	7.19	6.47	7.08	6.53	5.98	5.31
Month	Jun	Aug	Sep	Oct	Nov	Dec
Wind velocity(m/s)	4.23	5.05	5.52	5.25	5.75	6.00

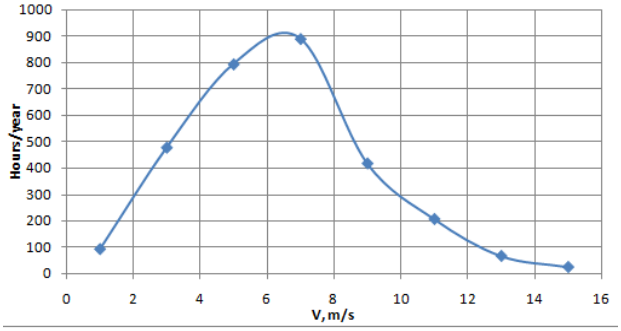


Figure 1 – Distribution of wind velocity over

Repetition of wind velocities is an important power characteristic. It shows how many times in a month or year the same velocity was. The graphical interpretation of wind velocity distribution is presented in Figure 1.

According to Lukutin et al. (2015), the wind power plant will be effective if the yearly average wind velocity is more than 4 m/s. The yearly average wind velocity in the Koyda village is 5.86 m/s.

3. Designing of wind-diesel power plant

Determination of the load consumers

The hourly load graph of the Koyda village with the highest energy needs in the year is presented in Figure 2

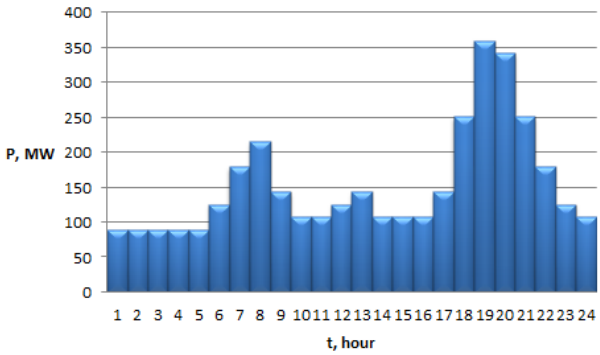


Figure 2 – Hourly load graph

Selection of wind turbine generator

For this hybrid power plant I selected two wind turbines "CONDOR AIR WES 380/50-60" with nominal power 60 kW and nominal wind velocity 9 m/s. The load coverage by wind turbines is presented in Figure 3.

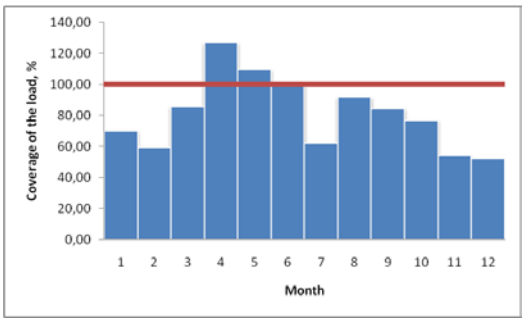


Figure 3 – Coverage of load by wind turbines

According to graph in Figure 3, two wind turbines cover 74% of year energy needs in the village or 527 280 kWh. The extra energy is only 2%. The fuel consumption reduced by 93.6 tons per year.

4. Configuration of wind-diesel power plant

The Koyda village already has diesel power plant, so, the wind turbine will operate with diesel generators in parallel. To perform such mode, it is necessary to improve the control algorithm for hybrid power plant by addition in the configuration next equipment: a multipurpose inverter, which can work as stand-alone and in parallel with network, synchronize devices. The configuration is presented in Figure 4.

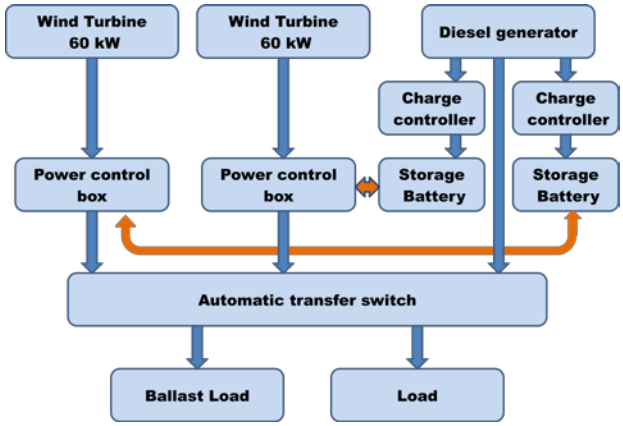


Figure 4 – Hybrid Power Plant with Inverter Diesel power installation [1]

The advantage of an inverter diesel power installation is reducing of fuel consumption during low load modes due to reducing of diesel generator rotation frequency.

5. Economic analysis

Nowadays, the Koyda village has power supply from diesel generators. To upgrade the system to wind-diesel power plant we need to buy and install next equipment: two wind turbines, two power control boxes, two charge controllers, two blocks of storage batteries and automatic transfer switch. The total cost of investment is 16 082 640 RUB.

Annual costs for fuel, repair and maintenance etc. with only diesel generators are 17.6 million RUB. With additional two wind turbines the costs will be dramatically decreased to 5.3 million RUB.

The next step is to calculate minimal price for both variants. The tax rate in Russia is 20%. To calculate discount rate I used the formula from the CAPM model [6]:

$$ER_i = R_f + \beta_i(ER_m - R_f)$$

Taking into account that average inflation is 7.32%, β is 0.544 for power sector, market risk premium of Russia is 9.43%, the discount rate is equal to 8.47%.

In calculation was used real price for electricity in the Arkhangelsk Region (5.65 RUB/kWh). NPV of the project is -44 335 734 RUB. NPV is negative and it can be explained that the Koyda village is supplied by state company, so people have subsidies and pay for electricity less than it costs.

NPV of the power plant if we will do nothing and power will be generated only by diesel generator is -228 211 058 RUB. It is even less than in previous case. Therefore, if the power plant will be upgraded with wind turbines, we can save 183 875 323 RUB.

According to Arkhangelsk Region Agency of tariffs and prices, price for 1 kWh in the Arkhangelsk Region is 5.65 RUB. To calculate minimal price it is necessary to make NPV equal to zero. Therefore, minimal price for 1 kWh in case of wind-diesel power plant is equal to 9.64 RUB. Therefore, to make the project at least lossless, we need subsidies in amount of (9.64 – 5.65) = 3.99 RUB for 1 kWh. Minimal price for 1 kWh in case of diesel power plant is 24.79 RUB.

Weather can be predicted but sometimes the expectation is not so accurate. The graph in Figure 5 shows how will change NPV with price 9.64 RUB for 1 kWh in case of lower load coverage by wind turbines.

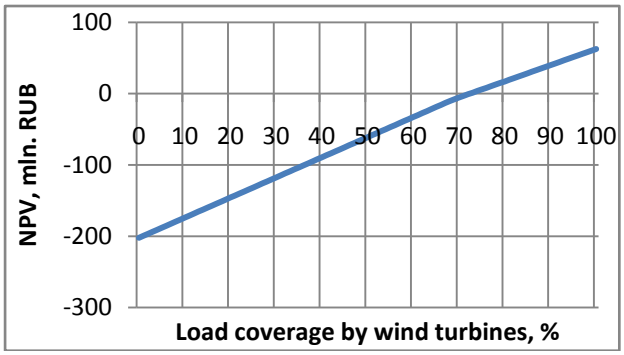


Figure 5 – Dependence of NPV on load coverage by wind turbines

The graph in Figure 5 has a linear trend. According to the graph, NPV becomes positive when wind turbines generate 74% of power that load consumes or more.

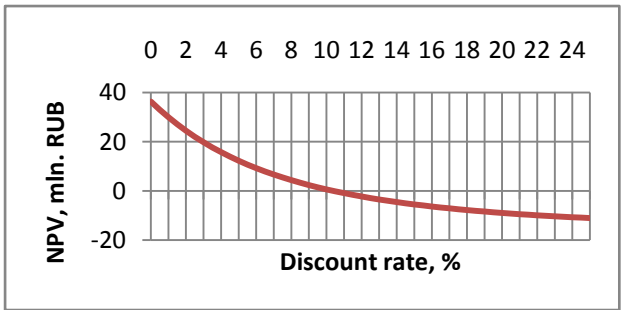


Figure 6 – Dependence of NPV on discount rate

The graph in Figure 6 has exponential trend. According to the graph, the greater discount rate the less NPV. NPV is positive when discount rate less than 8.47%.

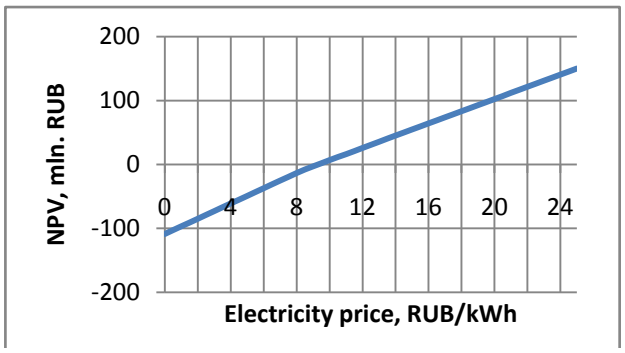


Figure 7 – Dependence of NPV on electricity price

The graph in Figure 7 has linear trend. According to the graph, the more electricity price the more NPV. NPV becomes positive when price is more than 9.64 RUB.

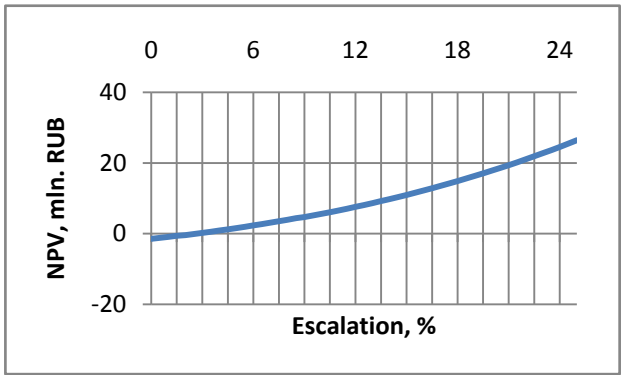


Figure 8 – Dependence of NPV on escalation

The graph in Figure 8 has exponential trend. According to the graph, the greater escalation the greater NPV. To make the project profitable it is possible to take escalation equal to 3% or more.

6. Conclusion

In this paper the wind-diesel power plant has been designed. The power plant based on diesel power plant in the Koyda village. The plant was upgraded with two wind turbines with power of 60 kW each. It help to economy 93.6 tons of fuel per year and it decreased cost of energy by 61%. The investment in the project is 16 082 640 rubbles and its NPV negative but we can save 183 875 323 RUB in 20 years in compare with diesel power plant.

References

[1] LUKUTIN, B.V., MURAVLEV, I.O., PLOTNIKOV, I.A. *Sistemi elektrosnabzheniya s vetrovimi i solnechnimi elektrostanciyami* [Power supply systems with wind and solar power plants]. Tomsk: Tomsk Polytechnic University, 2015.

[2] RenEn – Renewable energy, http://renen.ru/res-in-germany-generated-37-5-of-electricity-for-three-quarters-of-2017/ - Renewable energy in Germany, 2017, date of reference: 07.01.2019.

[3] KOMAREVTSEV, M.A., POPOV, A.N., YARKOV, D.A. Sistema Energoobespecheniya na baze gibridnoy ustanovki [Power Supply System Based on Hybrid Power Plant]. Lomonosov North Arctic Federal University, 2018.

[4] Solbat – solar batteries, http://www.solbat.su/meteorology/insolation/, solar insolation tables, date of reference: 07.01.2019.

[5] Weather timetable, https://rp5.ru/Архив_погоды_в_Абрамовском_Маяке, weather record in Abramovskiy lighthouse, date of reference: 07.07.2019.

[6] BREALEY, R.A., MYERS, S.C., ALLEN, F. Principles of Corporate Finance, 10th ed. McGraw-Hill/Irwin, 2010.

About the Author

Andrey Kurganovskiy was born in 1995 in Russia. He graduated from Tomsk Polytechnic University with a Bachelor's degree "Electrical Power Engineering". In 2017 he started his Master's degree in "Electric Power Generation and Transportation" in Tomsk Polytechnic University. Through a joint double-degree program he is doing his Master's degree in "Economics and Management of Electrical Engineering" at Czech Technical University in Prague.

Evaluation of the effectiveness of photovoltaic modules

Mikhail IGNATENKO

Department of Economics, Management and Humanities, Czech Technical University in Prague,

Technická 2, 166 27 Prague, Czech Republic

ignatmik@fel.cvut.cz

Abstract. *The use of renewable energy sources is rapid now due to limited traditional fossil fuels. Solar energy is an interesting and perspective area of study in power sector because of its affordability and relative cheapness. Of course, the point is to collect the energy and convert it into electricity. Solar panels are created to make it possible. For the optimal choice of solar panels, it is necessary to evaluate the required characteristics, as well as to know how and under what conditions they were obtained. In the technical description of solar panels, the manufacturer indicates the parameters obtained in the laboratory. Often, technical description of panels, which is obtained in laboratory, does not correspond to the data when panels are operated in real conditions. The purpose of my work is to evaluate the efficiency of photovoltaic modules in real-life conditions. To accomplish the aim, I will conduct experiments with the Aiyima solar panels.*

Keywords

Renewable energy source, photovoltaic module, solar panel, solar cell.

1. Introduction

Nowadays, renewable energy is one of the most promising and fast-growing energy industries. The problem of the lack of fossil resources at the moment is not acute, but its depletion is inevitable. Moreover, progressive humanity is engaged in the search for alternatives where renewable energy is predominant. The most widespread technology in this field is hydropower plants. However, in regions where there are no large rivers and the wind speed is not high enough, the most feasible type is solar energy. The sun is a huge, inexhaustible and safe source of energy, equally owned and accessible to everyone.

The chosen topic is relevant because solar panels are increasingly used not only in industry but also in private use, and the consumer often does not see accurate data about the product, but only approximate results, obtained in the laboratory. In the course of work, I will try to estimate how much the parameters declared by the manufacturer differ from the real ones.

There are two main types of photovoltaic modules: monocrystalline and polycrystalline panels on silicon base.

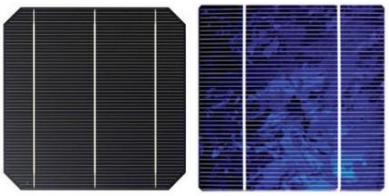


Figure 1. Monocrystalline (left) and polycrystalline (right) solar cells [1]

The manufacture of solar cells based on monocrystalline silicon allows to obtain the highest efficiency about 19–22% [2]. However, topic of my work is based on research of qualities of polycrystalline type of panels. The main advantage of polycrystalline solar cells is that they are cheaper (around 6%), since the cost of the material is lower, but efficiency is also lower (around 15%), lower space efficiency, and a shorter lifespan since they are more affected by hot temperatures [3]. It is optimal to use this type of panels if there is no goal to obtain maximum power generation per unit of installed capacity and if there is no significant difference in illumination levels for a long period.

For evaluating effectiveness of solar modules, experimental model of solar panel based on polycrystalline solar cells was made. It has low power, but it is enough to make analysis about its expediency. The design provides the ability to change the angle of the panel to the solar radiation and to change load.

2. Author's experiment

In order to evaluate effectiveness of solar panels in conditions of real work, I had conducted an experiment. The idea is to test the parameters of solar cells and compare results of testing with its passport parameters.

Characteristics of Aiyima solar cell:

- Material: Polycrystalline Silicon
- Size: 52x52 mm
- Power: 0.43 W
- Voltage: 0.5 V
- Current: 0.86 A
- Efficiency: 17%

The experimental model is a supporting structure for eight solar cells welded together and a separate block of resistors to change the load.

Supporting structure is a varnished wooden desk with four bolt feet, mechanism for changing and fixing the angle (six working positions) and four output banana type plugs.

Solar cells was soldering consequently. This scheme involves connecting the “plus” of the first cell with the “minus” of the second, and the output of external wires from the “minus” of the first cell and the “plus” of the last.

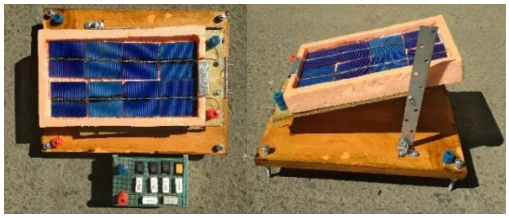


Figure 2. Author’s constructed experimental model (views from different angles)

Because of the fact that panel was constructed by me in not precise conditions, values of angles are not standard and there is even negative value of angle in the first case. Table 1 shows obtained data.

Experimental data was collected under the following conditions:

- Date / Time: 16.07.2018 / 12:07;
- Ambient temperature: 22° C;
- Temperature of the panels before experiments: 35° C;
- Temperature of the panels after experiments: 46° C;
- Weather: sunny, cloudless.

Table 1 – Experimental data

R=0 Ohm			R=384 Ohm			R=22.3 Ohm		
α, deg	I, mA	U. V	α, deg	I. mA	U. V	α. deg	I. mA	U. V
-2.27	0	4.55	-2.27	10.8	4.53	-2.27	181	4.3
1.98	-	4.38	1.98	11	4.35	1.98	172	4.2
8.97	-	4.51	8.97	10.8	4.5	8.97	180	4.26
14.77	-	4.52	14.77	11	4.5	14.77	175	4.2
18.80	-	4.43	18.8	11.1	4.41	18.8	170	4.21
39.32	-	4.45	39.32	11	4.39	39.32	170	4.19

Table 2 – Experimental data (continuation)

R=9.8 Ohm			R=8 Ohm		
α. deg	I. mA	U. V	α. deg	I. mA	U. V
-2.27	420	3.9	-2.27	500	3.13
1.98	401	3.85	1.98	465	3.32
8.97	425	3.8	8.97	519	3.45
14.77	400	3.81	14.77	491	3.4
18.8	420	3.78	18.8	515	3.53
39.32	412	3.8	39.32	490	3.62

After obtaining data, VAC graph is built. Figure 3 shows VAC of first and last tilt angles of model. Its form corresponds to standard VAC of PV panels.

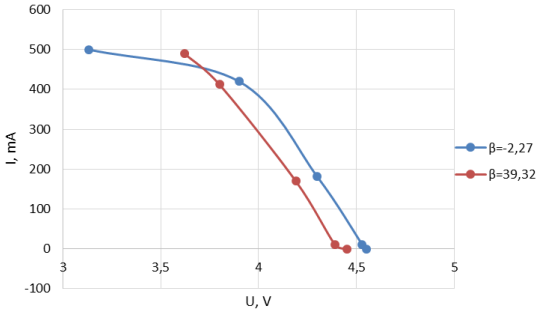


Figure 3. Volt-ampere characteristics (VAC) of experimental installation for different tilt angles β

3. Calculation of basic parameters of effectiveness

I was conducting experiments with six different tilt angles. The first value of β=-2.27° will not be taken into account during calculation of optimal angle and solar radiation, because it cannot be optimal.

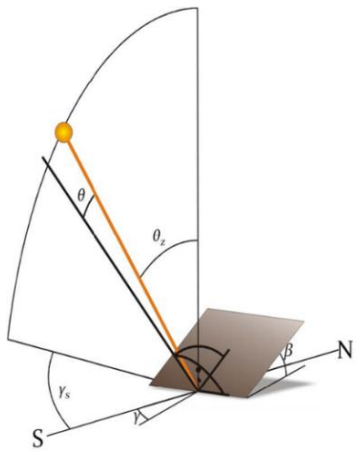


Figure 4. Tilted plane with tilt angle β, plane azimuth angle γ and incidence angle θ, additionally solar zenith angle θz, solar azimuth angle γs [4]

First step is to determine the optimal incidence angle of the beam radiation on it. The incidence angle depends on the following parameters: geographical location (latitude and longitude), time of the day and of the year and plane orientation (tilt angle and azimuth angle).

Following preliminary calculations, graph of optimal angle was made in Mathcad. As it could be seen from the Figure 5, the biggest value of solar radiation is observed at the angle of 39.32°.

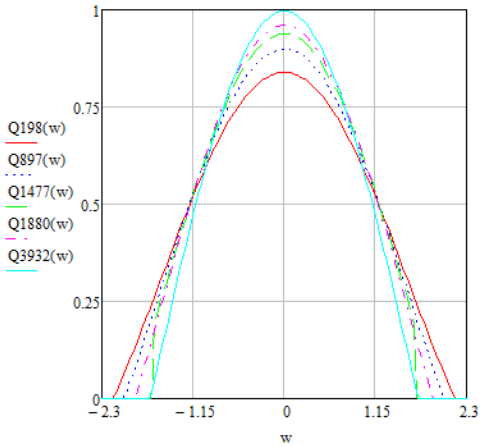


Figure 5. Optimal tilt angle of the panel

After calculations of basic parameters of effectiveness, it is possible to compare results with the data from [5].

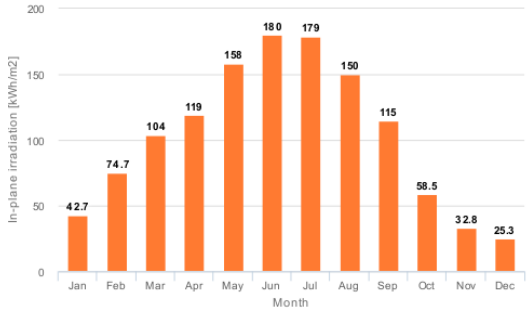


Figure 6. Monthly in-plane irradiation

Result is different from the calculated variant because of the several reasons. First of all, unknown value of the losses in calculations. Second reason is that time of experiment is approximate because conducting experiment requires some extra time for switching resistors etc.

4. Calculation of the efficiency of solar cells and solar panel

When all necessary data is obtained, I will calculate efficiency of solar cells and solar panel used in experiment. In addition, I will compare results with the data provided by the manufacturer.

Insolation in July:

$$\delta = 179 \ kWh/m^2; \text{ [Figure 4]}$$

Area of one solar cell:

$$S_C = 52 \cdot 52 = 2704 \ mm^2;$$

$$S_C = 2704 / 1000000 = 0.002704 \ m^2;$$

Area of a solar panel consisting of 8 solar cells:

$$S_p = 0.002704 \cdot 8 = 0.021632 \ m^2;$$

Energy of one solar cell at 100% efficiency:

$$E_C = \delta \cdot S_C;$$

$$E_C = 179 \cdot 0.002704 = 0.484 \ kWh;$$

Energy of solar panel at 100% efficiency:

$$E_C = \delta \cdot S_p;$$

$$E_p = 179 \cdot 0.021632 = 3.872 \ kWh;$$

Power of one solar cell at 100% efficiency:

$$P_C = \frac{E_C}{N \cdot h};$$

$$P_C = \frac{0.484}{31 \cdot 7} = 0.00223 \ kW;$$

where,

N – Number of days in July;

h – Daylight hours.

Power of solar panel at 100% efficiency:

$$P_p = \frac{E_p}{N \cdot h};$$

$$P_p = \frac{3.872}{31 \cdot 7} = 0.0178 \ kW;$$

Rated nominal power of one solar cell:

$$P_{NC} = U \cdot I;$$

$$P_{NC} = 0.5 \cdot 0.86 = 0.00043 \ kW;$$

Hence, efficiency of one solar cell in the selected area:

$$\eta_c = \frac{P_{NC} \cdot 100\%}{P_C};$$

$$\eta_c = \frac{0.00223 \cdot 100\%}{0.00043} = 19.28\%;$$

Real power of panel according to experimental data:

$$P_{RP} = 490 \cdot 3.62 = 0.00177 \ kW;$$

Efficiency of solar panel in the selected area:

$$\eta_p = \frac{0.00177 \cdot 100\%}{0.0178} = 9.94\%.$$

As we can see, obtained results are differ from the data declared by the manufacturer (19.28% and 9.94% against 17%). The difference in efficiency is significant. Moreover, as experiment conducted in July in cloudless weather conditions, we can assume that in winter season or in cloudy weather, results would be even worse. I see several reasons of difference between obtained data and efficiency declared by manufacturer:

1) Dirty surface of the panel. During the process of constructing experimental model (soldering, placing), solar cells were slightly polluted.
2) Temperature of the cells. With increasing of temperature, the efficiency of solar cells decreases. Higher temperature increases the flow of electrons, which causes an increase in current and voltage drop. The voltage drop is higher than the increase in current. Therefore, the total power is reduced, which leads to the fact that the module operates with less efficiency [6]. In our case, temperature of the panel before and after experiment is 35° and 46° relatively.
3) Optimal angle according to [5] is 43°, however design of experimental model provides angle of 39.32°.
4) Soldering of solar cells and contacts was carried out manually, which could lead to additional losses.

5. Efficiency from the economic point of view

All solar panels have some efficiency. The higher efficiency, the less amount of solar panels required to obtain necessary value of energy. That is, due to the increase in efficiency, the rated power increases with the same or smaller dimensions of a panel.

The most common methods for increasing the efficiency of panels are:

1) Solar tracking system. This mechanism rotates panels, changing the angle and direction in order to maximize receiving of insolation. Sometimes it could be used for cleaning from snow, but also there is a separate snow removal system.
2) Snow removal system. When snow covers surface of the panel, amount of generated energy is reduced. The automatic heating system of solar modules with the help of sensors receives information that the panel is covered with snow, and begins the process of heating. Melting snow will create conditions for the sliding of the snow layer under the action of gravity.

There are various types of these systems from simple static to complex dynamical. However, before using them it is necessary to compare its impact on the effectiveness and its price.

As an example, I will make a model of a photovoltaic power plant for Pervopashensk village in Tomsk region. In modeling, I will show the impact of including a solar tracking system into photovoltaic power plant.

Diesel generators supply Pervopashensk village with the tariff = 34.75 RUB/kWh [7]. Investor decide to build photovoltaic power plant in order to cover some part of consumers energy demand. Diesel generators will cover remaining part.

Table 2 – Photovoltaic power plant data

Without trackers		With trackers	
Parameter	Value	Parameter	Value
Cost of solar panels	3 444 800 RUB	Cost of tracking systems	2 776 700 RUB
Efficiency	17.2%	Average increased efficiency	27.9%
Amount of panels	400	Amount of tracking systems	25
Annual energy output	134 550 kWh	Annual energy output	175 560 kWh
Annual profit	2 691 000 RUB	Anuual profit	3 511 200 RUB

Assumptions made:

- Income tax = 20%;
- Discount rate = 16%;
- Cost of installation work = 10% from cost of the system;
- Selling price = 20 RUB/kWh;
- Lifetime of panels = 20 years.

In order to calculate NPV of installing tracking systems with perform of following calculations.

Since I compare two models, the annual profit is difference between annual profit of the variant with solar tracking system and without it.

Annual payment:

$$PMT = 3\,511\,200 - 2\,691\,000 = 820\,200 \ \text{RUB};$$

Present value annuity factor:

$$PVAF = \frac{Investments}{PMT};$$

Investments are presented as a sum of solar tracking system cost and the cost of its installation work:

$$Investments = 2\ 776\ 700 + 277\ 670 = 3\ 054\ 370;$$

$$PVAF = \frac{3\ 054370}{820200} = 4.7;$$

$$PVAF = \frac{1-(1+r)^{-T}}{r};$$

$$4.7 = \frac{1-(1+0.16)^{-T}}{0.16};$$

Payback period:

$$T = 9.4 \text{ years};$$

NPV for 20 years:

$$NPV = 835\ 898 \text{ RUB};$$

There are various types of solar tracking systems in the market. Chosen type is one of the cheapest because it designed for 16 solar panels, that is why it is profitable to implement it. Projects payback period is less than 10 years and its NPV for the lifetime of solar panels is equal to 835 898 RUB. I also considered case with 8 solar tracking systems designed for 50 panels with price of 1 450 000 apiece. In this case, investment on tracking system is huge and it is not profitable to install it.

6. Conclusion

In the course of the work, experiments were performed with polycrystalline solar modules of the Aiyima Company. Optimal angle of inclination of the panel for the selected area was calculated. The obtained data allowed me to compare the efficiency, declared by the manufacturer, with the efficiency of work in real conditions. As expected, the resulting efficiency is significantly (by 7.06% lower) different from the stated. A model of installation of systems that increase the efficiency of the panels was also built. Calculations performed show that it is profitable to implement chosen type of the system.

References

[1] Tipy solnechnykh paneley (Types of solar panels). Available at: https://www.atmosfera.ua/pvsolar/tipy-solnechnyx-panelej/. Accessed on 19.11.2018.

[2] Testirovanie parametrov solnechnykh batarey (Testing of Solar Panels Parameters). Available at: www.solarhome.ru/pv/buyer_guide/ptc_conditions.htm. Accessed on 25.10.2018.

[3] Seven Different Types of Solar Panels Explained. Available at: www.greenmatch.co.uk/blog/2015/09/types-of-solar-panels. Accessed on 19.11.2018.

[4] MATTHIAS GUNTHER – *Advanced CSP Teaching Materials. Chapter 2. Solar Radiation*, - DLR (German Aerospace Center).

[5] PHOTOVOLTAIC GEOGRAPHICAL INFORMATION SYSTEM Available at: re.jrc.ec.europa.eu/pvg_tools/en/tools.html. Accessed on 04.01.2019.

[6] Vliyanie temperatury na effektivnost' solnechnykh batarey (Influence of temperature on solar cells efficiency). Available at: https://gws-energy.ru/blog/40-vliyanie-temperatury-na-effektivnost-solnechnykh-batarej/. Accessed on 14.04.2019.

[7] Vozobnovlyaemye istochniki energii Tomskoy oblasti (Renewable energy sources of the Tomsk region) Available at: http://green.tsu.ru/tomres/?p=117. Accessed on 28.04.2019.

About the Author

Mikhail IGNATENKO was born in 1996 in Kazakhstan. In 2017, he graduated from Tomsk Polytechnic University with a Bachelor's degree in "Electrical and Power Engineering". In 2017 he started the Double Degree program in "Electric Power Generation and Transportation" in Tomsk Polytechnic University and "Economics and Management of Electrical Engineering" in Czech Technical University in Prague.

Benefits of Battery Storage for Charging Stations

Ing. Lukáš Dvořáček

Dept. of Economics, Management and Humanities, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

dvoral14@fel.cvut.cz

Abstract.

Following the EU 2030 targets, the reduction of GHG emissions becomes ever more pressing. All stakeholders will have to seek for solutions to reach the reduction of GHG emissions across sectors - not only in the electricity and heat generation, but also in the public and especially private transport sector. Mainly the field of new vehicles should actively change. The European Parliament specified that 40% of new cars to be zero- or low-emission vehicles by 2030 targets.
With the increase of electric vehicles (EVs), it will be important to find suitable points of charging with regards to transport infrastructure and complete the existing system of energy network with superchargers. The existing literature and research have so far mainly focused on the location of the charging stations in the primary transport nodes like gas stations along the highways and roads among the cities and have rarely considered the fact that in the first stage, most EV will be used as city vehicles. It is highly likely, that due to the parameters of current batteries and small distance which the EVs are able to achieve, the EVs will initially be used for urban and suburban traffic.

Keywords

Battery storage, fast-charger station, electric vehicles, electric taxis, charging site, consumption diagram, 2030 targets

1. Introduction

The key parameter, which needs to be taken into account when allocating the charging sites in the existing energy infrastructure, is the fundamental difference between the charging an electric car and refueling a car with a combustion engine.

The home charging stations created a new way the users of electric vehicles are able to recharge their cars very comfortably. This option has to be considered before designing a network of recharging stations [1]. One of the most significant benefits of home charging of electric vehicles will be a lower price per kWh compared to a business fast-charging station [2], together with the higher degree of comfort and less time spent charging in public charging sites [1], [2]. On the other hand, the disadvantage of home charging solution is the necessary upfront investment in the home charger set and sure the investment in the own parking place or garage [3].

However, not all users (especially in urban areas) have their garage or parking space. Most of the population in cities live in flats without a private garage or a private parking space [3]. These larger groups of users leave their cars in public parking places nearby apartments. They do not usually have an option to buy a home charging set and use it for charging electric cars.

The problem of charging electric vehicles outside the home can be divided into two main points. The first point is charging a large number of electric cars during a long time with low energy power and the second point is charging electric vehicles with fast charging stations in the shortest possible time.

Design of charging stations, for instance, in the large parking spaces nearby apartment buildings could resolve the issue of recharging electric vehicles for many users living in this urban area. In this large urban parking places it is expected that the user park their car and let it charge for a longer period, usually more than 6 hours. During this time and among a lot of electric vehicles recharging, it is possible to control that network stability is not compromised and the energy capacity of connection this area is fully used.

Completing the charging site in this urban area with a battery could be a backup for potentially increasing the number of charging stations. But the expected amount of simultaneously charged electric vehicles is not possible to use only the charging site with a battery to increase in charging station energy power or a reduction in charging time without further interfering with the infrastructure of the superior network. A better place to build a charging point supplemented with a large capacity battery is the city charging site with a large number of public fast charger stations or areas close to parking places for electric taxis. There we can expect high demand for fast recharge

In view of the efforts of all European cities to expand electric vehicles and electric taxi cars [4], the EU will be building new super-charging stations in the city centers especially at the places reserved for the taxi cars.

It can be assumed that the first wave of charging sites in the urban areas will be created by adding fast-charging stations to the existing gas stations because existing gas

stations usually have a perfect position in the city in terms of transport infrastructure. Also, it is expected that this upgrade of gas stations by new charging stations will mean lower investment than building new electric charging site.

From an energy perspective, these important areas can have a problem with the capacity of the energy networks, and it may not be possible to build the charging site with the regard of the current stability of energy network.

However, in the fast charging stations, where short charging time is a priority (as for instance those for taxi vehicles), the method of controlling the delay of charging is not permitted. Also, it may not be possible to add another power line to the city place, which in turn could stop this project of building enough volume of powerful charging stations. In such case, the shared battery storage is the only solution how to get a charging site which satisfies energy consumption of electric vehicles without needing additional infrastructure investment.

2. Charging site installation

The present model assumes a location in an urban area that is connected to a 640 kW power transformer station. The charging station is not the only entity that is connected to the transformer station; the other connection is a housing unit with 360 apartments. The model assumes that all flats are connected to a central heating system and thus do not have their water heater and are not equipped with any other source of heat like an electric boiler or set of connectors. It means that all apartments have a D02 or D01 tariff with a single-tariff electricity rate and flats have a circuit breaker 26A.

With this assumption, it is possible to extend the calculation by a type delivery diagram from 4 sections (TDD4). The delivery diagram characterizes the yearly consumption of electricity at the customer and is also used to evaluate deviations. For calculation, the model uses an annual consumption diagram whose measured values also take into account the temperature development in the Czech Republic during the year [5].

Averages for each of the 24 hours of the day were calculated from the month measured values of the delivery diagram. To estimate the average value of the electric power, which the whole house unit required in each day, it is necessary to multiply these values by the difficulty factor, the number of apartments and amper value of circuit breakers. Difficulty factor (1) takes into account the behavior of individual users and the possibility of switching on especially powerful appliances in individual apartments at the same time.

$$K_n = K_\infty + \frac{(1 - K_\infty)}{\sqrt{n}} \quad (1)$$

Where K_n is the difficulty factor; K_∞ is the contemporaneity coefficient for infinite number of flats and n is number of flats [6]

Important for the charging site is that the difference between the reserve power for the housing unit and the maximum power level of the transformer station (TS) can be used to charge the battery.

It is important to consider the impact of battery charging on the consumption diagram. If the battery is charged by maximum power from the transformer station, it will reduce the charging time. The downside of the battery charging will be the unwanted fluctuations in the consumption diagram. It will be particularly recognizable at intervals with low energy demand (Figure 1, 2).

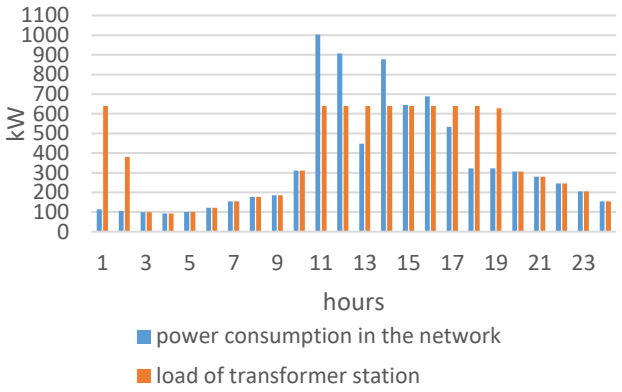


Figure 1 utilization of transformer station power reserve from 100%

Figure 1 shows a short time needed to charge the battery and the power peak that the charging in the consumption diagram in May caused.

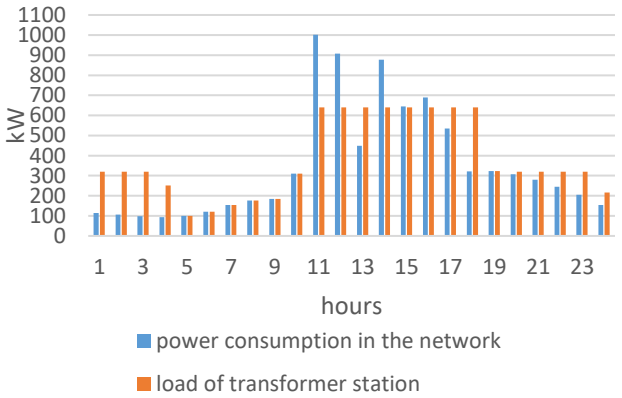


Figure2 utilization of transformer station power reserve from 50%

Figure 2 shows the consumption diagram when charging the battery with half the charging power of the transformer station than in Figure 1. This way of charging also involves unwanted fluctuations in the consumption diagram in May.

Therefore, the model considers controlled battery charging and charges the battery out of hours with high demand for energy to smooth out small variations in consumption. In the case of uncontrolled battery charging, the model achieves a faster battery charge but causes unwanted fluctuations in the consumption diagram. For this

reason, the model consumes constant energy from the transformer station throughout the time interval with a small demand for electric energy. The benefits of constant energy consumption are also smoothing of smaller fluctuations in the consumption diagram at time intervals beyond the main demand to a constant value. The stored energy from the intervals with less energy demand is used to smooth the power peaks in the load diagram when the energy demand is higher. Very often these peaks are caused by charging a large amount of EV at the same time.

2.1 Model of a charging site

2.1.1 Technical part

The model is considering building a charging point with shared battery storage exclusively for electric taxis. The model is based on a study of the charging frequency of electric taxis at one particular urban charging site in Singapore but with more frequent arrivals of electric taxis [7]. From the used values, 20 electric taxis were simultaneously charged during the main quarter-hour.

Another input condition of the model from [7] is the fixed working hours of taxi service. The first taxi cars are leaving at 6 am and ending at 6 pm. During this time interval they can use the charging site to charge their cars. The charging station is not used after working hours.

During the working interval it is assumed that the electric taxis arrive to the charging stations with almost empty batteries and after 45-minute recharging, they depart with a fully charged batteries. To meet the 45-minute limit for almost full battery charge and to consider the average battery capacity used in the electric car, it is necessary to use a 50 kW DC or higher power charging station for charging. It was found out from the real measurement that a 50kW DC charger connected to a three-phase distribution system will charge a car battery for 10kWh in 12 minutes and 30kWh in another 24 minutes [8].

Supply Type / Charger Rating	AC/DC	Rated power*	Time to charge 10kWh
3 phase, DC / Rapid	DC	50kW	12 mins
3 phase, DC / Rapid	DC	120kW	5 mins
3 phase, 32A per phase / Fast	AC	22 kW	27 mins
3 phase, 60A per phase / Rapid	AC	43kW	14 mins

**Simplistically (*ignoring Power *Factor correction)*

Tab. 1. Supply types and charging times. *[8]*

Therefore, the model assumes that in 45 minutes the charging station will charge 30 kWh of electric power. Considering the actual energy value that will be stored in the electric car battery after 45 minutes, it will be necessary to consider system-wide losses due to heating, transforming the AC voltage to DC and others. These losses range from 15% to 20% during charging [9].

However, for the implemented model, the priority is the value of the energy that has been released from the battery or the power network. The model knows from data collected by tracking the arrival of individual taxi cars to the charging point at 15-minute intervals that the largest consume value of energy was 250 kWh (Figure 3).

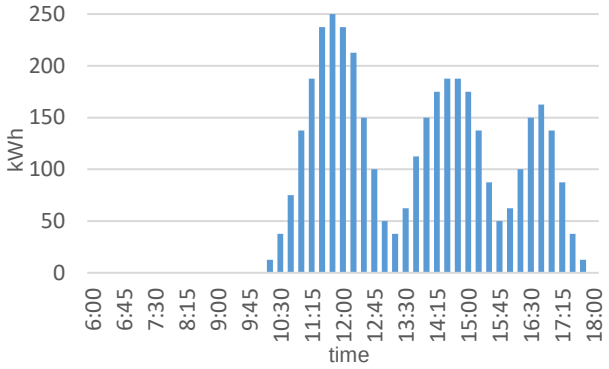


Figure 3 Consumption of EVs in the 15-minutes intervals. [7]

The installation of the charging site in the urban area increases the demand for electric power especially in the primary charging hour of taxi vehicles. The increase in consumption is mainly due to the charging of electric vehicles at the same time. However, due to the same start of working hours and similar mileage for each taxi cannot be eliminated (Figure 4).

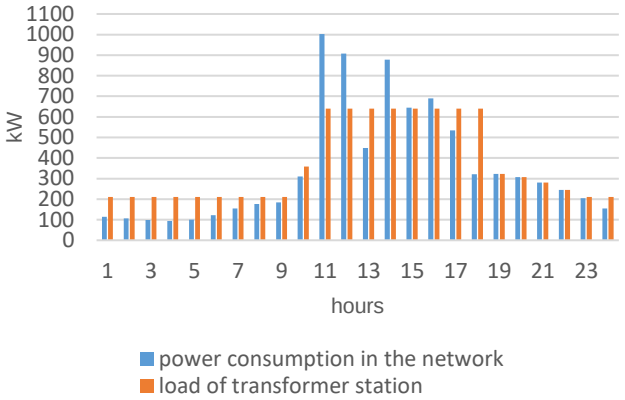


Figure 4 consumption diagram for controlled battery charging

The implementation of this charging site without the use of a battery required a power boost not only for the local transformer station but also increase of all power elements towards the power distribution network. In the urban area are usually some parts of the power grid, where increasing the necessary power elements in the network can't be realized. Limiting factors are generally power limits, current throughput, network stability or high investment to strengthen the entire power network.

Compared to increasing the power limits of all elements of the power grid, the realization of charging site with a large capacity battery has a lot of benefits. One of the

advantages is the possibility of battery installation in places where the power increase of the transformer station is not possible due to the stability of the whole network.

Furthermore, the possibility of installing a higher quantity of charging stations at a given location (charging site) and thus obtaining the possibility of charging a plurality of electric vehicles at a set time without endangering the stability of the power network.

In terms of network stability, the battery located near to the point of large energy consumptions improves smoothing power peaks in the consumption diagram. It helps to stabilize the transmission network and, of course, to a stable frequency in the power grid.

2.1.2 Economic part

Apart from technical aspects, the economic factors need to be taken into account when constructing a charging point. One of the key issues will always be the place of realization. As already mentioned from the perspective of transport infrastructure and accessibility, the already-built gas stations are the ideal place to build a charging site. However, the gas stations may not have a sufficient capacity of energy connection. In particular, the requirements for the energy power of the local installed transformer station and the current throughput are usually unsuitable.

If these technical requirements for increasing the power of the local transformer station and all the energy elements can be realized towards the distribution network, it is also necessary to take into account all these costs related to the strengthening of the whole energy infrastructure.

Very often, the addition of a high-capacity battery charging station is a simpler and more cost-effective option, which does not require additional investments to strengthen the surrounding infrastructure although the initial investment in battery storage that exceeds 80,000 euros [10] is not negligible and may seem disadvantageous.

Significant but economically uneasy consideration will also be the consideration of charging site realization that could not be realized in the required location without battery storage.

Conclusion

The issue of the construction of recharging points with the expected growth of electric cars will be necessary and up to date. There will be an increase in the demand for energy consumption at the points of implementation of these stations and it will be necessary to take steps to ensure the stability of the entire power system.

One way to build a charging site at the required location, which does not endanger the stability of the entire power grid, is to strengthen the local transformer stations and all the energy elements towards the distribution network. However, often this method is not feasible from a technical or from an economic point of view.

The second option is to use the battery storage, which is a backup source to reduce the energy fluctuations in the energy demand during the highest charging station load. The advantage of this solution is that it does not require another change in the energy network and helps to smooth the fluctuation in the consumption diagram, as is also apparent from the model.

The implemented model presented a possible solution to the construction of the charging station by using battery storage. For real implementation of the model, it is necessary to improve the prediction of the difference between the current household consumption and the expected consumption of electric vehicles primarily. With more data obtained, it will be possible to determine better the required battery size and other costs necessary to implement the charging station.

References

[1] D. Lopez-Behar, M. Tran, T. Froese, J. R. Mayaud, O. E. Herrera, and W. Merida, “Charging infrastructure for electric vehicles in Multi-Unit Residential Buildings: Mapping feedbacks and policy recommendations,” Energy Policy, vol. 126, no. May 2018, pp. 444–451, 2019.

[2] “Cost of Charging an Electric Car”, Pod Point’. [Online]. Available: https://pod-point.com/guides/driver/cost-of-charging-electric-car. [Accessed: 5-Mar.-2019].

[3] Z. Wei, Y. Li, Y. Zhang, and L. Cai, “Intelligent parking garage EV charging scheduling considering battery charging characteristic,” IEEE Trans. Ind. Electron., vol. 65, no. 3, pp. 2806–2816, 2018.

[4] “Podnikatelé se mohou těšit na vyšší podporu elektromobility ze strany vlády: Podpora infrastruktury”, Byznys&energie. [Online]. Available: https://www.byznys-energie.cz/autor/jan. [Accessed: 1-Mar.-2019].

[5] Normalizované TDD’, OTE, a.s. [Online]. Available: https://www.ote-cr.cz/cs/statistika/typove-diagramy-dodavek-elektriny/normalizovane-tdd. [Accessed: 18-Mar-2019].

[6] ČSN 33 2130: Elektrické instalace nízkého napětí - Vnitřní elektrické rozvody, 3rd ed. 2015.

[7] H. Wang, R. L. Cheu, and E. Balal, “Operations of electric taxis to serve advance reservations by trip chaining: Sensitivity analysis on network size, customer demand and number of charging stations,” Int. J. Transp. Sci. Technol., vol. 5, no. 2, pp. 47–59, 2016.

[8] “SPIRIT Power change: EV Charging Knowledge Bank Understanding electric car charging”. [Online]. Available: https://www.spiritenergy.co.uk/kb-ev-understanding-electric-car-charging. [Accessed: 1-Mar.-2019]

[9] E. Apostolaki-Iosifidou, P. Codani, and W. Kempton, “Measurement of power loss during electric vehicle charging and discharging,” Energy, vol. 127, pp. 730–742, 2017.

[10] “Česko má třetí velké bateriové úložiště s kapacitou 1,3 MWh: Smart City”, IOT-network newa, 2018. [Online]. Available: https://cz.iot-nn.com/blog/2018/08/24/cesko-ma-treti-velke-bateriove-uloziste-s-kapacitou-13-mwh/. [Accessed: 18-Mar.-2019].

About Author

Lukáš Dvořáček - Ph.D. student in the Department of Economics Management and Humanities, FEE CTU in Prague

Prediction of household consumption based on the model

Ing. Martin Horák

Department of Economic, Management, and Humanities, Czech Technical University, Jugoslávských partyzánů 3, 166 27 Praha, Czech Republic

horakm11@fel.cvut.cz

Abstract. *In the electric power grid, an influence of intermittent sources starts not to be negligible. Increasing the percentage of production from renewable energy sources may increase demands for power regulation in the network. However, the power regulation is now mostly ensured through the control of fossil power plants. Exiting from the use of fossil fuels will further reduce the possibility of regulation. One of the solutions can be regulation on the customer side where they can actively manage their consumption. This solution requires the development of infrastructures such as smart meter and communication systems, which have been, however, uneconomic so far.*

Until the cost of technology is reduced, the way to develop a decentralized energy grid is by using the predictions of user consumption. Such estimate must be much more accurate than the currently available load profile of distribution tariffs. The present research, therefore, evaluates the changes in diagrams of households` electricity consumption based on usage of the different household electric appliances, prediction of lifestyle patterns, improving thermal and the technical properties of the buildings leading to a model to estimate the future consumption profile of households. All research is carried out on data from the Czech Republic.

Keywords

Energy load profile, Smart city, zero-energy-building, decentralization

1. Introduction:

With the use of new types or more efficient appliances and the installation of decentralized power sources, e.g., PV and (micro)cogeneration, the households are no longer just consumers and become so-called "prosumers" (producer and consumers). They produce electricity on their own or sell the surpluses return to the energy grid [1] In the electric power grid, an influence of intermittent sources starts not to be negligible. To achieving the RES EU target of 2030, the RES share for electrical generation will increase from this year's 13% to 18 % in 2030 and further increase according to vision EU 2050 [2] [3] [4]

However, electricity as a commodity differs from others in the difficulty of storing a large amount of energy for a long time. For short-term balance in the grid, the active management of a power plant system has to be used. At the time, the power inflowing to the grid from power plants must always correspond to the amount of energy outflowed, consumed, in the grid. If there is an imbalance in the network, the frequency in the grid is changing, later the overload and blackout may occur.

The original consumption pattern that is used for current electric profile prediction follows only the use of household's amenities like a water heater and central electrical boiler. This model has been so far sufficient for group hour load profile but not for single household prediction.

The consumption may vary due to minor influences like the behavior of householders or so-called rebound effect. This effect describes the fact that low consumption of new appliances can lead to more often use, and thus the total household consumption can increase. However, two more challenges in future load profile changes are; growing interest in electromobility that can cause the short-term high consumption electricity demand from the grid that distribution network may not be able to satisfy, and the increasing number of PV panels whose an intermittent production cause abnormality in consumer load profile.

The solution to this problem may be implementing battery repositories (or other technologies) so that the maximum energy produced in small decentralized units (PVs) is consumed at the area of production and do not increase the distribution grid load. Thanks to cost-effective technologies for storing energy (i.e., battery storage, domestic water heating, storage tanks for heating), deviations could be eliminated at the level of buildings, groups or small district. To properly design these local repositories, it is necessary to create a sufficiently accurate consumption pattern for individual buildings or units.

This project tries to evaluate the changes in diagrams of households' electricity consumption due to using another household electric appliances, changing lifestyle. The results of the analysis will fit into the draft of the updated load profile .

2. Statistical survey

To estimation future changes in consumption, the current state of household equipment and consumer behavior has to be analyzed. For this purpose, the latest statistical survey ENERGO 2015 [5] gathered between 7 July 2015 and 15 January 2016. The sample contains statistical data about 20,000 households, i.e., 0.5% of all permanently inhabited dwellings.

The survey is divided into several modules mapping the energy demands of the population.

- Housing module:
  Energy performance of the selected apartment/house, residential area, number of floors, year of construction, a way of house insulation, and other additional questions.
- Fuel module:
  For heating, water heating, cooking, lighting, cooling
- Lighting module:
  The structure of the light sources.
- Appliance module
  Only appliances with the most significant share of the total household electricity consumption were included (refrigerators, freezers, automatic washing machines, clothes dryers, dishwashers, televisions, computers, and air conditioning). Their energy class, age, and frequency of use.
- Transport module
  Use of cars and their characteristic parameters

At the end of the questionnaire, participants were asked about the total household income for further analysis.

To estimate further changes in household amenities, the values from the ENERGO 2015 are compared with the Odyssee-Mure statistic data for neighboring countries. During the ten years from 2005 to 2015, the consumption of electricity by households fell in the whole EU-28 by 0.9 % in the Czech Republic by 2.3%; the seventh largest decline in the EU-28 (Figure 1). [6]

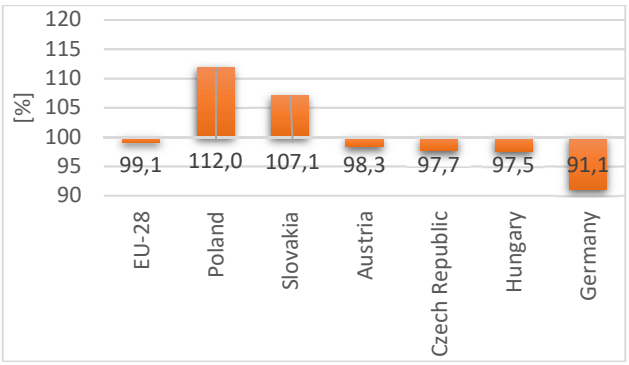


Figure 1 Changes in households electricity consumption between the years 2005-2015 *[6]*

A significant influence on electricity consumption has household electrical appliances growth. In households, there is a growing number of new types of appliances with specific consumption, such as entertainment technology, dryers, smart home appliances, etc.

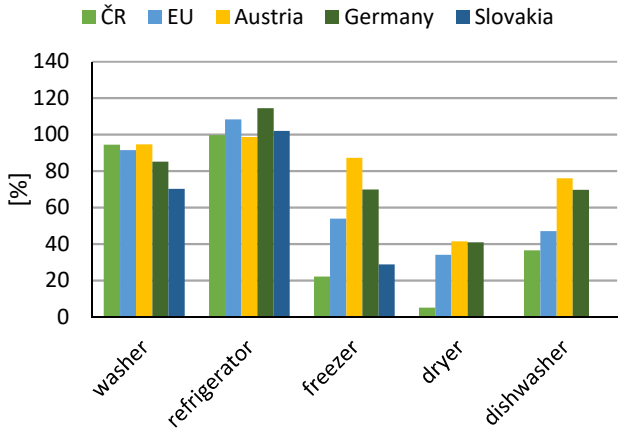


Figure *2* Home appliances *[6]*

Figure 2 shows the coverage of household appliances in the Czech Republic. It can be assumed that the coverage will approach the level of EU and Austria and Germany. Moreover, the natural replacement of the appliances will increase the average cost-effectiveness of the main appliances. A lifetime of appliances depends on many factors, but according to ENERGO statistic data in TTable 2 and the proportion of appliance age in Table 1 indicates that approx. 50% of refrigerators, washing machines, and dishwashers older than 6 years will be exchanged for Class A+, or better devices and consumption can decrease by about 30%. [5] [7]

Energy class (%)	A+++	A++	A+	A	B-D
refrigerator combined	8,9	20	25,6	33,2	12,3
refrigerator	6,3	13,7	19,6	32,9	27,5
freezer	5,3	12,6	19,3	33,8	29
dryer	23,4	32,3	23,8	14,8	5,7
dishwasher	11,9	26	27,4	27,3	5,4
washer	8,7	22,2	26,6	30,8	11,7

Table 1 Share of energy classes in household appliances owned by households in year 2015 *[5]*

Age (%)	2 yrs	2 -5 yrs	6-10 yrs	more
refrigerator combined	6,3	36,3	37,1	18,3
refrigerator	4,9	28,5	36,9	29,7
freezer	2,8	25,6	34,2	34,4
dryer	25,2	55,7	16,5	2,6
dishwasher	10,5	51,3	31,1	7,1
washer	9	40,5	36,5	14

Table 2 Average age of the appliances in households in year 2015 *[5]*

3. Smart metering data analysis

The smart meters data has been used for consumption estimation — the data contained ten annual load profile from households within five different distribution tariffs. Data from the smart meter contained consumption value for every 15 minutes and a timestamp for each entry. All the data comes from the region of Central Bohemia, namely the distribution power grid under the management of the company PREdistribuce, a.s.

This analysis aims to determine the predictability of consumption based on real household data. The data are processed regarding the assumption that the activities of the inhabitants related to the annual cycle and the natural weekly period.

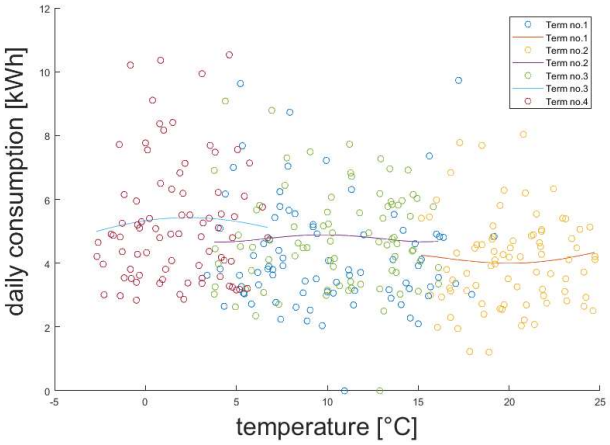


Figure 3 Temperature correlation

The annual consumption period is mostly associated with weather, ambient temperature. As an example, consumption during the summer months can be higher as usual by thanks to working air conditioning. On the contrary, in winter the consumption might be higher due to higher lighting requirements. For the annual period reason, the data has been correlated with ambient temperature in the year as shown in Figure 3.

The weekly period is mainly influenced by the householder's daily routine. Therefore, the analysis focuses on the similarity of the daily consumption diagrams during the same weekday. Since it cannot be eliminated the sessional dependency of consumption, the calculation was always created within a shorter period than a year, usually one month. Figure 4 shows the mean consumption and standard variation for March Monday of 2017.

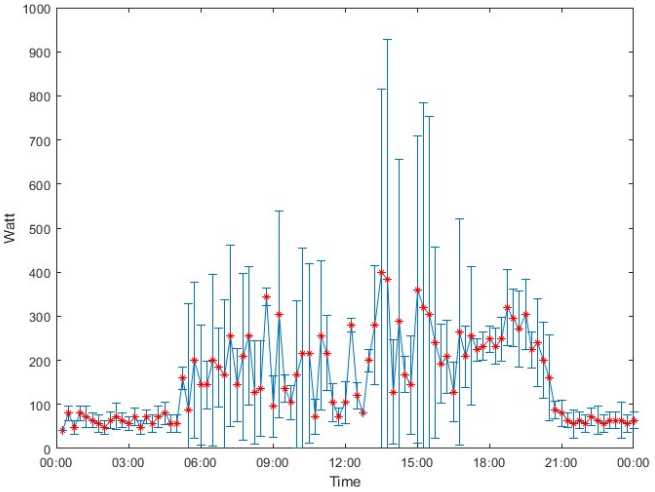


Figure 4 Five-week Monday consumption

4. Discussion

The research uses both statistical data (ENERGO 2015), and real data from already installed smart meters. The survey shows that nearly half of household appliances have an energy class worse than the minimum value that can be sold in the EU after 2013. Assuming the average replacement of appliances is about 7 - 10 years, it can be expected that approximately half of the major appliance will be exchanged for class A+ or better devices and consumption from these appliances decrease approximately about 20%. Also, due to using of new facilities in households, we can expect significant changes in consumption pattern after 2020.

The pilot analysis of the real-time data from households shows the correlation among consumption, ambient temperature, and periodic living pattern. Analyzed data, for instance, showed that a family with the parent on maternity leave consumes about 70% of daily energy during the lunchtime unlike households with residents with a full-time job whose consumption is moved to afternoon or evening hour. Still, the identification of individual smaller appliances contribution in total consumption profile remains a challenge for further research.

With the incoming entry ban for diesel vehicles into European city centers, partial replacement by electric cars can be expected. Assuming an average annual mileage of Prague driver is about 8 thousand kilometers [8], and average consumption of an electric car is 15 kWh / 100 km [9], average household consumption can grow by a non-negligible 2.7 MWh per year. This amount can be a part of household consumption in case of household in a house. However, as it is not possible to charge the car in flats, in those cases the direct consumption would not be affected.

5. Conclusion:

Developments in the consumption of electrical appliances cause changes in the total consumption of households. Together with the effort to reduce energy losses in flats and houses, household dependence on electricity supplies is growing. Moreover, with the EU regulation on the minimum standard for newly built houses from 2020, greater deployment of heat pumps, recuperation unit, and smart heating system deployed can be expected. When comparing household amenities in neighboring countries, we can also expect the use of dryers or dishwashers that are not so common in the Czech Republic.

Acknowledgments:

Thanks to prof. Ing. Jaroslav Knápek, CSc. for help and suggestions for this work. This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS18/136/OHK5/2T/13.

Bibliography

[1] "Egú Brno," 20 09 2018. [Online]. Available: https://www.egubrno.cz/wp-content/uploads/2018/10/TZ_Energetika_2018.pdf. [Accessed 23 03 2019].

[2] E. COMMISSION, "Policy 2030," [Online]. Available: http://data.consilium.europa.eu/doc/document/ST-169-2014-INIT/cs/pdf.. [Accessed 13 3 2018].

[3] E. COMMISSION, "Roadmap2050," 2011. [Online]. Available: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52011DC0885&from=EN.. [Accessed 13 3 2018].

[4] M. o. o. a. trade, "Státní energetická koncepce," [Online]. Available: https://www.mpo.cz/dokument158059.html. [Accessed 23-Mar-2019].

[5] Č. s. ú. |. ČSÚ, "Energo 2015," 12 2014. [Online]. Available: https://www.czso.cz/csu/czso/energo-2015.. [Accessed 13 3 2018].

[6] Odyssee-mure, "Energy Efficiency Trends and Policies in Buildings," 2015. [Online]. Available: http://www.odyssee-mure.eu/publications/br/energy-efficiency-in-buildings.html. [Accessed 13 03 2019].

[7] S. A. E. B. Anette Michel, "Energy efficiency of White Goods in Europe:monitoring the market with sales data – Final report. ADEME," 2016.

[8] Odyssee-mure, "distance travelled by car," 2015. [Online]. Available: http://www.odyssee-mure.eu/publications/efficiency-by-sector/transport/distance-travelled-by-car.html. [Accessed 2019].

[9] U. D. o. E. -. E. E. a. R. Energy, "fueleconomy.gov," [Online]. Available: https://www.fueleconomy.gov/feg/PowerSearch.do?action=noform&path=1&year1=1984&year2=2019&vtype=Electric. [Accessed 20 03 2019].

About the Author

Martin HORÁK
was born in 1993. In 2017 graduated (Ing.) at the Department of Telecommunications Engineering on the Faculty of Electrical Engineering at Czech Technical University in Prague. He received a master's degree in electric power engineering on subject developing the RFID measurement system for Factory 4.0. Currently studying, a Ph.D. in the Department of Economics, Management and Humanities on the Faculty of Electrical Engineering at Czech Technical University in Prague. His scientific research is mainly focused on research of smart grids and telecommunications.

"Comparative analysis of micro cogeneration for family house in the Czech Republic and the Russian Federation"

Viktoriia PEREBOEVA

[1] Dept. of Economics, Management and Humanities, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

perebvik@fel.cvut.cz

Abstract. *Cogeneration technology (in other words, combined heat and power production technology, CHP) is widely used for centralized heating and electricity supply since XX century, but now the small-sized installations for little amount of consumer become more and more popular. Micro cogeneration units include cogeneration units with a maximum capacity below 50 kW. The advantages of using micro cogeneration are considered in the article. The research of the same object (family house) in different conditions allows making comparative analysis for two chosen countries from economical point of view. The most influential factors are the different climate conditions, fuel costs and government's support measures.*

Keywords

Micro cogeneration, combined heat and power production, comparative analysis, decentralized energy

1. Expediency of micro cogeneration

The main feature of cogeneration is production two forms of useful energy from one source. Thereby cogeneration maximizes use of primary fuel energy. Also, one of defining elements of cogeneration is the closeness of the energy conversion unit to a customer.

- Efficiency Benefits

Cogeneration offers energy savings ranging between 15-40% when compared against the supply of electricity and heat from conventional power stations and boilers due to co-production of heat and electricity.

Because cogeneration units are close to consumers, it also avoids transmission and distribution losses of electrical energy from central power system.

- Environmental Benefits

By utilizing of heat that would otherwise be wasted during the production of electricity, CHP systems require less fuel to produce the same amount of energy. Thereby, greenhouse gas emissions, such as carbon dioxide (CO2), as well as other air pollutants like nitrogen oxides (NOx) and sulfur dioxide (SO2), are reduced.

- Economic Benefits

Reduced energy costs: CHP has high efficiency. Also, CHP systems typically use natural gas which is often cheaper than purchased electricity.

Avoided capital costs: CHP can often reduce the cost of replacing heating equipment.

In addition, a CHP system can be configured to operate on a variety of fuel types, such as natural gas, biogas, coal, and biomass; therefore, a facility could be undependable from the high fuel prices.

- Reliability benefits

CHP can provide high-quality electricity and thermal energy independently of power grid and its outages.

2. Design of micro cogeneration for family house

Residential buildings, especially huge houses, located out of city, often have a need for high ambient temperatures during heat season and have high demand for hot water. For comparative analysis was chosen detached family house for 4 persons. The house was built according to the standard project, which can be realized as in the Czech Republic as in the Russian Federation. . The house location is Dobřejovice village in the Czech Republic and Tomsk in the Russian Federation. Living area of house is 217,7 m^2, overall size of the building -1080 m^3. Both localities are supplied by natural gas.

2.1 Collating of electrical and thermal load data

The difference in the thermal load data is based on different average outside temperature during heating period.

Locality	Dobřejovice (1 variant)	Tomsk (2 variant)
Average outside temperature during heating period	-12 °C	-40 °C
Heat losses of house	11,647 kW_t	15,98 kW_t
Annual energy requirement for heating and HWS	37,6 MWh	51,7 MWh

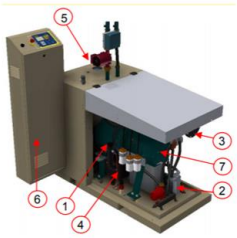
Tab. 1. *The main thermal load data for chosen localities*

The maximum electrical load is **5,6 kW$_e$**, the total annual electrical energy consumption is **17,535 MWh**.

2.2 Selecting of micro cogeneration unit

Based on gained electricity and heat consumption data is possible to choose appropriate cogeneration unit. One of the biggest providers is TEDOM Group Company. TEDOM Group provides selling and service of cogeneration plants around the world, including the Czech Republic (ČEZ Energo s.r.o.) and the Russian Federation (TEDOM RU).

For both localities was chosen CHP unit «MICRO 7» is working on natural gas. The electrical output is 6,5 kWe and heat output is 16 kWe. The unit has size 1300x1120x1380 mm and weight 645 kg. Lifetime is 15 years.



1. Generator
2. Plate heat exchanger
3. Heat exchanger for product of combustion
4. The oil tank
5. Connection points
6. Electrical switchboard
7. The internal combustion engine

Fig. 1. *Microgeneration unit and its main components*

3. Main economic factors

3.1 Fuel costs

The initial investments for both variants will be almost equal, because of one provider. The more important parts of costs are costs of fuel. On the one side, Russia has lower gas price, but, on the other side, the annual consumption of fuel is less for Czech Republic.

Locality	Dobřejovice (1 variant)	Tomsk (2 variant)
Price of natural gas for 1 m^3, CZK / RUB	12,9	5,12
Total annual gas consumption, m^3	4098	5 470,8
Constant monthly payment, CZK/month	318,26	-
Total costs, CZK / RUB	57 444	28 010

Tab. 2. *The fuel costs for both variants*

3.2 Electricity cost

Other important factor is electricity cost. The differences between costs of own production of electricity and price of electricity from the grid determines the benefit of cogeneration project.

Locality	Dobřejovice (1 variant)	Tomsk (2 variant)
Price of electricity for 1 MWh, CZK / RUB	2,764	3,5
Total annual electricity consumption, MWh	17 535	
Constant monthly payment, CZK/month	53	-
Total costs, CZK / RUB	49 102	61 373

Tab. 3. *The electricity costs for both variants*

3.3 Government's support measures

In the Czech Republic there is annual green bonus for electricity from a CHP plant. This bonus depends on installed capacity of unit and operating hours. In the Russian Federation there is no government support of cogeneration.

For CHP plant with installed capacity below 200 kW$_e$ and operation hours more than 3000 the support is 1283 CZK per 1 MWh.

3.4 The selling of electricity to the grid

The main feature of own production of electricity from the economical point of view is opportunity to sell surpluses of electricity to the grid. The prices and rules of such selling are regulated by laws. For units less than 10 kW of power the simplified procedure is used and no license required.

4. Economic evaluation

4.1 Cogeneration mode

There are different ways to produce electricity and heat by cogeneration unit. In this case, production must cover all electricity demand. The produced power must cover monthly peak loads. The output is regulated by rated power of generator and cannot be less than 50%. Heat and electricity production after a valid regulation is presented on figures 2 and 3.

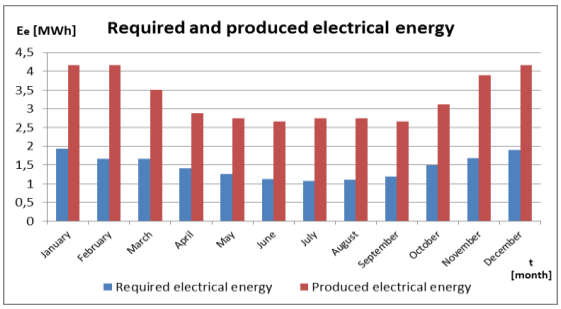


Fig. 2. *Required and produced electrical energy*

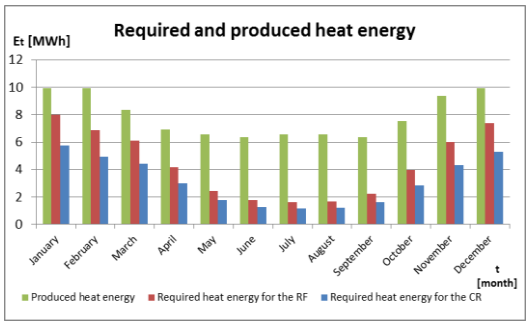


Fig. 3. *Required and produced heat energy*

In the table 4 are presented investment and variable cost for chosen cogeneration mode.

Locality	Dobřejovice (1 variant)	Tomsk (2 variant)
Price of cogeneration unit [CZK / RUB]	500 000	560 000
Repairs and maintenance [CZK/RUB per year]	9 715	29 145
Price of natural gas for 1 m^3 [CZK / RUB]	11,83	5,12
Total annual gas consumption [m^3]	15 823	
Constant monthly payment [CZK/1000 m3]	247 056	-
Total costs of gas [CZK / RUB]	4 096 353	81 014

Tab. 4. *Investment and variable cost for chosen cogeneration mode*

The necessary condition to get the green bonus for units less than 1 MW is proof of useful using of primary energy. Since, in this variant there are big losses of heat, it is not possible to pretend to the green bonus.

4.2 Calculation of Net Present value

Nominal discount rate is equal risk-free rate since this project is not investment project with making profit goal.

According information from national banks, for the Russian Federation risk-free rate is 8,6 % , for the Czech Republic risk-free rate is 1,82 %.

Inflation rate for the Czech Republic is 2% according National Bank forecast. For the Russian Federation inflation rate is 5,3 %.

Locality	Dobřejovice (1 variant)	Tomsk (2 variant)
Risk-free rate, r_f [%]	1,8	6,75
Inflation	3	5

Tab. 5. *The data for calculation of discount rate*

There are two ways to account the inflation in the NPV calculation. The first is usage of nominal cash flows, which includes the effect of inflation with nominal discount rate. The second is usage of real cash flow with real discount rate, which includes the effect of inflation. In this paper will be used first method.

Net present value:

$$NPV = \sum_{1}^{T} \frac{CF_t}{1+r_n} - inv \qquad (1)$$

where NPV - Net present value, CF_t – nominal cash flow at t-year, *inv* – initial investments, r_n - nominal discount rate.

$$CF_t = CF_{t\,real} \cdot (1+inf)^t \qquad (2)$$

CFt – nominal cash flow at t-year, CFt real – real cash flow at t-year, *inf* - inflation rate.

For comparison were calculated NPV for initial variant and cogeneration mode. This calculation includes the selling to the grid the excess of electricity at price 1,01 RUB for the RF and 1 CZK for the CR per 1 kWh.

NPV / Locality	Dobřejovice (1 variant)	Tomsk (2 variant)
Initial variant [mln. CZK / mln.RUB]	-1,986	-2, 018
Installation of cogeneration unit [mln. CZK / mln.RUB]	-3,244	-1,834

Tab. 6. *The Net Present Values for all variants*

4.3 Sensitivity analysis

For sensitivity analysis dependencies of NPV from discount rate for both countries were built.

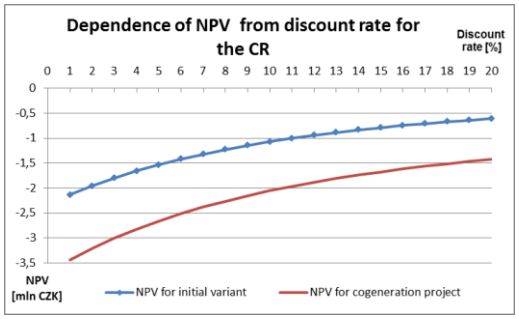


Fig. 4. *Dependence of NPV from discount rate for the Czech Republic*

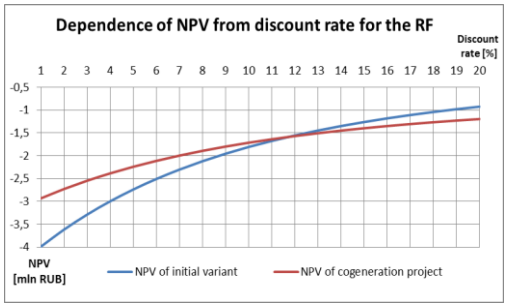


Fig. 5. *Dependence of NPV from discount rate for the Russian Federation*

As can be seen from figures, at low discount rate the NPV has low value. With grown of discount rate NPV grows at accelerated pace. For the Czech Republic initial variant better than cogeneration. But for the Russian Federation, at the discount rate less than 11%, cogeneration becomes more profitable.

The calculations show that cogeneration for family house is not so much profitable today. A few reasons of this and problems of domestic cogeneration are presented below.

- Low heat demand. The one of the most problem of domestic cogeneration is excess heat recovery. This problem can be solved, if family house have almost uniform annual heat demand, for example, swimming pool.
- Wrong mode of cogeneration was selected. The full covering of electricity demand not obligatory condition. In case of uniform heat demand, we can cover it by cogeneration and take lacking electricity power from the grid.
- Limited regulation options. The family house has low heat and electricity demand, so the cogeneration unit operates on half of rated power, but output is still not such low, as it needed. The market of cogeneration units cannot offer units, suitable for family house conditions yet.
- High costs for investments. The costs of standard connection to the electricity grid and heat equipment were not included in the calculations, but, obviously, cost of cogeneration plant for today many times more.
- Low cost of selling exceeds of electricity. In case of excess of electricity there is opportunity to sell it and make some profit, to cover initial investments. But, the cost of it cannot be more than wholesale price that is it very low. Moreover, the procedure for processing necessary documents takes a lot of time and makes additional costs.

Finally, based on calculated NPV, the conclusions can be done. The NPV of cogeneration project for the Czech Republic is in times 1,63 worse less than initial variant. It is caused by extremely high price of gas. But, for the Russian Federation cogeneration is better than initial variant in 1,1 times. In spite of the absence of government support, cogeneration for the Russian Federation is more perspective because of greater heat demand and less fuel price.

5. Acknowledgements

Research described in the paper was supervised by Ing. Bc. Blanka Kucerkova, FEL CTU in Prague.

References

[1] M.PEHNT, M.CAMES, C.FISCHER, B PRAETORIUS, L.SCHNEIDER, K. SCHUMACHER, J.VOSS. – *Micro Cogeneration: Towards Decentralization Energy System*, Springer. 2010, ISBN-13: 9783642064982

[2] KOVARO GROUP. *Proekt odnoetazhnogo doma v Prage.* [Online] Available from: http://www.ksreality.ru/mainmenu-nedvizhimost-v-chekhii/doma-v-prage/138 [Accessed 6 December 2018].

[3] OTE. *Electricity Load Profiles* [Online] Available from: https://www.ote-cr.cz/en/statistics/electricity-load-profiles/normalized-lp?date=2018-12-16 [Accessed 16 December 2018].

[4] TZBINFO. *Venkovní výpočtové teploty a otopná období dle lokalit* [Online] Available from: https://vytapeni.tzb-info.cz/tabulky-a-vypocty/25-venkovni-vypoctove-teploty-a-otopna-obdobi-dle-lokalit [Accessed 29 December 2018].

[5] TEDOM. *Overview of produced models* [Online] Available https://www.tedom.com/en//chp-units/natural-gas/ [Accessed 10 March 2019].

[6] Brealey, Myers, Allen. Principles of Corporate Finance, McGraw-Hill, 2016, 12th edition, ISBN:9781259253331

[7] Damodaram online. *Data current* [online] Available http://people.stern.nyu.edu/adamodar/New_Home_Page/datacurrent.html [Accessed 12 April 2019].

About Author

Viktoriia PEREBOEVA was born in Russia in 1995. In 2013, she entered to the Tomsk Polytechnic University with field of study «Electrical Power Engineering». In 2017, she started Master's double-degree program implemented with Czech Technical University in Prague.

Finantial Time Series Forecasting With Wavelets

Jaroslav SCHÜRRER[1]

[1]Masaryk Institute of Advanced Studies, Czech Technical University, Kolejní 2637/2a, 160 00 Praha, Czech Republic

schurjar@cvut.cz

Abstract. *In this paper, we address the problem of how to model non-stationary financial time series by means of wavelets and use this model for comparison with the standard univariate autoregressive moving average model (ARIMA). Financial time series include multiscale features. That is, an observed time series may contain several structures, each occurring on a different timescale. We use Discrete Wavelet Transformation as a preprocessing tool to create a model which is consequently used for further analysis in our case prediction.*

Keywords

Wavelet transformation, ARMA.

1. Introduction

A time series is as 'the sequence of events which constitutes or is measured by time.' Time series are used to characterize the time course of the behavior of a wide variety of biological, physical and economic systems. Industrial production data, interest rates, inflation, stock prices, and unemployment rates, amongst other time-serial data, provide a measure of the health of an economy.

The interest in time-varying events ranges from gaining a better understanding of the underlying system producing the time series to being able to forecast the future evolution of the data generating process. Researchers have generally adopted time series analysis methods in an attempt to comprehend time series data. Such methods are based on the assumptions that one might discern regularity in the values of measured variables in an approximate sense, and that there are, many at times, patterns that persist over time.

Followers of the Efficient Market Hypothesis assert that price changes in financial markets are random and it is impossible to consistently outperform the market using publicly available information [6]. On the other hand, [5] argue that the Efficient Market Hypothesis is an economically unrealizable idealization that 'is not... well-defined and empirically refutable', and that the Random Walk Hypothesis is not equivalent to the Efficient Market Hypothesis.

It has also been argued that an informationally efficient market is impossible [3]), individuals exhibit bounded rationality [9], and market expectations may be irrational [4].

We think that financial markets are not always efficient and make sense to analyze time series o understand the past and predict the future. During time-series forecasting the historical values are collected and analyzed in order to develop a model describing the behavior of the series. When the time series is non-stationary, it is very difficult to identify a proper global model. To overcome this problem, an efficient way is to use the wavelet decomposition technique in the preprocessing step. The Wavelet transform provides a useful decomposition of time series, in terms of both time and frequency, permitting us to effectively diagnose the main frequency component and to extract abstract local information from the time series.

1.1. Discrete Wavelet Transform

The Discrete Wavelet Transform is a special case of the Continuous Wavelet Transformation that provides a compact representation of a signal in time and frequency that can be computed efficiently. The DWT is defined by the following equation:

$$W(j,k) = \sum_j \sum_k x(k) 2^{-\frac{j}{2}} \psi(2^{-j} n - k) \qquad (1)$$

where $\psi(t)$ is a time function with finite energy and fast decay called the mother wavelet.

The DWT analysis can be performed using a fast, pyramidal algorithm based on convolutions with quadrature mirror filters [7]. In the pyramidal algorithm the signal is analyzed at different frequency bands with different resolution by decomposing the signal into a coarse approximation and detail information. The coarse approximation is then further decomposed using the same wavelet decomposition step. This procedure can be repeated for further decomposition where at every level, the filtering and subsampling will result in half the number of samples thus half the time resolution and half the frequency resolution.

This is achieved by successive high pass and low pass filtering of the time domain signal and is defined by the following equations

$$y_{high}[k] = \sum_n x[n]g[2k-n] \tag{2}$$

$$y_{low}[k] = \sum_n x[n]h[2k-n] \tag{3}$$

where $y_{high}[k]$, $y_{low}[k]$ are outputs of the high pass and low pass filters (g represent high pass and h low pass filter) after subsampling by 2. Each subband contains half the samples of the neighboring higher frequency subband. Because of downsampling the number of wavelet coefficients is exactly same as the length of input signal. This procedure can be repeated for further decomposition where at every level, the filtering and subsampling will result in half the number of samples thus half the time resolution and half the frequency resolution. Following picture 1 schematically depicts this procedure.

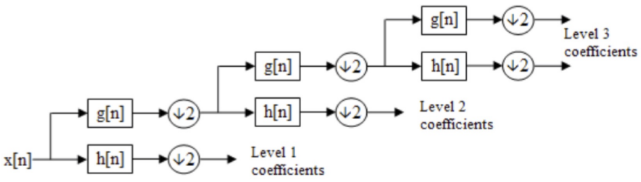


Fig. 1. Wavelet decomposition scheme.

Because of filtering the length of input signal plays important role. Wavelet algorithms expect the input length to be a power of two. If length is not suitable we have to use some extrapolation of the input data in order to extend the signal before computing the Discrete Wavelet Transform using the cascading filter banks algorithm. There are several methods of signal extrapolation that are mainly used: zero-padding, constant-padding, symmetric-padding and periodic- padding.

There exist different wavelet families, where Haar is simplest one followed by Daubechies (both orthogonal) and ends with biorthogonal wavelets.The extracted wavelet coefficients provide a compact representation that shows the energy distribution of the signal in time and frequency.

2. Autoregressive modeling

In this section, we report experiments conducted to investigate the effects of data preprocessing with Wavelets compared to univariate Autoregressive Integrated Moving Average model (ARIMA). The Autoregressive Integrated Moving Average (ARIMA) method models the next step in the sequence as a linear function of the observations and residual errors at prior time steps.

It combines both Autoregression (AR) and Moving Average (MA) models as well as a differencing pre-processing step of the sequence to make the sequence stationary, called integration (I). The notation for the model involves specifying the order for the AR(p), I(d), and MA(q) models as parameters to an ARIMA function, e.g. ARIMA(p, d, q). An ARIMA model can also be used to develop AR, MA, and ARMA models.

Figure 2 depicts financial time series preprocessing with wavelets. There are three steps which are used to produce prediction:

- DWT of original data
- Forecast of DWT coefficients at level one
- IDWT of forecasted DWT coefficients

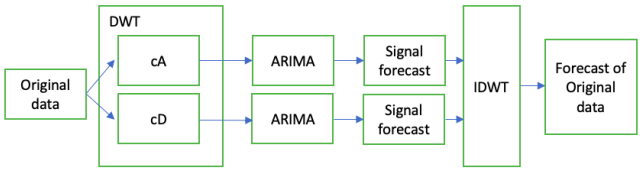


Fig. 2. Wavelet based approach for financial time series prediction

Firstly we decompose original data with DWT to obtain detailed $cD[n]$ and approximation $cA[n]$ coefficients at level one. During DWT we can use several wavelets (Haar, DB2, Coiflets, Biorthogonal ...). Next we forecast one data point with ARIMA method to model the next step in the sequence as a linear function of the observations and residual errors at prior time steps for both detailed and approximation coefficients to get $cA[n+1]$ and $cD[n+1]$. To obtain optimal order for ARIMA model we firstly evaluate an ARIMA model and then secondly we evaluate for different sets of order parameters (p,d,q), and choose the best one according to the mean squared error (MSE)At last we use IDWT to the new detailed and approximation vectors to get forecast of original signal.

Note that in the wavelet-based approach we can use instead of ARIMA another suitable forecasting method, for example ARMA, neural networks and so on. As we mention above different wavelet families can be used to decompose original financial time series into cA and cD coefficients.

For financial time series with large oscillations the D4 wavelet coefficients have larger values than those reached using the Haar wavelet, which means that they captures the high-frequency oscillations of the financial time series better than the Haar wavelet. That is the reason why we choose Harr and DB4 wavelets.

2.1. Real Word Time Series

The real-world dataset used in our analysis comprises three technology shares (Apple, Google and Microsoft) and the NASDAQ Composite which is a stock market index of the common stocks and similar securities listed on the NASDAQ stock market. NASDAQ is one of the three most-followed indices in US stock markets. Each series records the daily closing price in time interval starting at 1.1.2007 and ending 30.3.2019. Figure 3 depicts financial time series used in analysis.

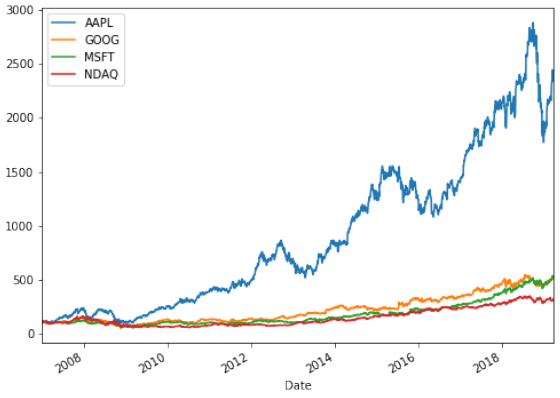


Fig. 3. Stock price of Google, Microsoft, Apple and NASDAQ.

The analysis will use a time interval starting on 1/1/2007 until 30/3/2019. We do not directly analyse the original or logarithmic price series, but we use log-returns calculated as follows:

$$r_t = \log(\frac{S_t}{S_t - 1}) \tag{4}$$

2.2. Experiments

For experiments we choose Daubechies D2, Daubechies D4, Coiflet 2 and Symlet 2 wavelets. In every case we predict one data point forward into the future. This procedure can be used to obtain several one-step predictions.

	Apple	Google	Microsoft	NASDAQ
ARIMA	**190.45**	**1172.70**	**117.90**	**87.46**
DB2	**189.49**	**1169.49**	**117.81**	**87.01**
DB4	**188.50**	**1169.49**	**118.35**	**87.48**
Coif2	**188.71**	**1169.49**	**116.96**	**86.68**
Sym2	**189.49**	**1169.49**	**117.82**	**87.02**

Tab. 1. Autoregresive modeling of selected financial time series.

We can see differences in prediction by ARIMA and wavelet preprocessing depicted in table 1. Differences in calculated values reflect different wavelets used during processing and prediction versus ARIMA model.

3. Conventions and notations

ARMA	Autoregressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
cA	Approximation coefficients
cD	Detail coefficients
DWT	Discrete wavelet transformation
IDWT	Inverse DWT
MSE	Mean Squared Error

4. Conclusion

This paper presented financial time series forecasting with wavelets. We demonstrate wavelet-based approach for financial time series prediction. This approach uses DWT to decompose original financial time series to detailed and approximation coefficients which are than forecasted and transformed back with IDWT to get forecasted original financial time series.

Acknowledgements

Research described in the paper was supervised by Prof. M. Vošvrda, Institute of Information Theory and Automation of the ASCR.

References

[1] DAUBECHIES, I. The Wavelet Transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 1990, vol. 36, no. 5.

[2] GENCAY, R., SELCUK, F, WHITCHER, B. An Introduction to Wavelet and Other Filtering Methods in Finance and Economics. San Diego: Academic Press, 2002

[3] GROSSMANN, S., STIGLITZ, J. On the Impossibility of Informationally Efficient Markets. In *American Economic Review, 70: 393–408.*

[4] HUBERMAN,G., REGEV, T. 'Contagious Speculation and a Cure for Cancer: A Nonevent That Made Stock Prices Soar. In *Journal of Finance, 56: 387–396.*

[5] LO, A. W., MACKINLAY, A. C. A Non-Random Walk Down Wall Street. Princeton: Princeton University Press, 1999.

[6] MALKIEL, B. A Random Walk Down Wall Street, 8th ed. New York: W. W. Norton & Co., 2003

[7] MALLAT, S. G. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. In *IEEE. Transactions on Pattern Analysis and Machine Intelligence, Vol.11,1989,674-693.*

[8] PERCIVAL, D.B., WALDEN, A.T. *Wavelet Methods for Time Series Analysis*, 1 ed. New York: Cambridge University Press, 2006

[9] SIMON, H. Models of Bounded Rationality. Cambridge, Mass: MIT Press, 1997

About Authors...

Jaroslav SCHÜRRER is a student of combined study course in Computational Economics and Finances on Masaryk Institute of Advanced Studies, Czech Technical University.

Polyaniline electric conductivity in the presence of ammonia molecules

Hana ŠUSTKOVÁ[1]

[1]Dept. of Microelectronics, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

sustkhan@fel.cvut.cz

Abstract. *Polyaniline as a conducting polymer is proposed to react with gas molecules and thus change its electrical conductivity. Therefore, in this paper a polyaniline molecule, in form of emeraldine salt, and its electrical conductivity in or without presence of NH_3 were studied by means of numerical modelling (QuantumATK). It was shown, that the IV characteristic of PANI variety emeraldine salt molecule reaches a saturation current of $14.1\mu A$ at 0.25 V. In the presence of the studied gas, NH_3, the electrical conductivity tends up to copy the IV characteristic of pure PANI without gas and saturates at similar voltage but reaches lower currents. The saturation currentr are 12.3 - 12.5 μA. This value is 12 - 13 % lower than the saturation current of pure PANI. These values are measurable - therefore, PANI in the form of emeraldine salt should be applicable for a NH_3 gas sensor.*

Keywords

Polyaniline, electric conductivity, IV characteristics, gas sensor, modelling.

1. Introduction

Polyaniline - a conducting polymer

Polyaniline, a well-known conducting polymer, seems to be interesting not only for future sensor applications. Being mechanically and chemically a polymer, polyaniline (PANI) can approach even electrical conductivity like metals. Thanks to π-conjugated bindings, electrical conductance is possible and can be adjusted by doping of the NH-groups. And not only its possibly high electric conductivity, the low-cost fabrication or environmental stability are important for future usage as well. PANI can be prepared either chemically or electrochemically, in many variations and with different electrical properties [1].

For gas sensor research, the emeraldine salt, a variation of PANI, is in focus. The basic PANI variety, leucoemeraldin, is a clear polymer, and namely an electrical insulator, like the second PANI variety, the blue polymer emeraldine (in the base form). Next, there is also a semi-conducting violet variety of PANI, pernigraniline. Putting a defect into the chain of insulating/low conducting PANI, one or mote polarons may appear. Thus by means of proton doping (of emeraldine base) or electron reducing (leucoemeraldine base) of every second nitrogen atom - there will appear 2 polarons per one 4-benzene-unit - the green emeraldine salt can be produced from emeraldine base or leucoemeraldine base, reaching a conductivity of 400 S/m [3].

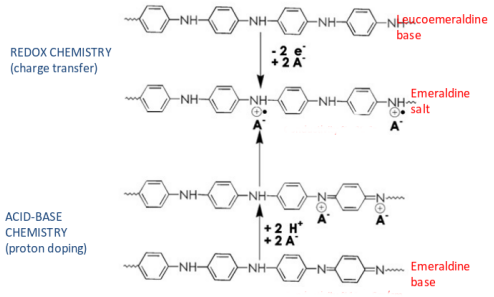


Fig. 1. Polyaniline varieties and their fabrication (from [3])

According to previous research, the PANI in the form of emeraldine salt may react with a group of gas molecules, among others with NH_3 or NO. However, the change in the electric conductivity of this polymer, while exposed to these gases, may be relatively low [2]. In this paper, the electrical conductivity of PANI with and without gas molecules in the nearest area of the polymer will be studied in order to find out the usability of polyaniline as a gas sensor.

2. Numerical approach

Computation in this paper was done by QuantumATK, using semi-empirical tight-binding methods. During benchmark, the Extended Hückel Method [4] was chosen. Device systems and their IV characteristics were then computed using non-equilibrium Green's functions formalism (NEGF), used in the Extended Hückel Method as well [5].

Extended Hückel method is based on the Hückel method but considering not only π-orbitals but also σ-orbitals. In this method, the electron-electron repulsions are

not included explicitely; the total energy comes from the sum of every electron in the molecule [4].

As for the Extended Hückel Method being used for bulk and molecule, self-consistent and non-self-consistent part of the tight-binding Hamiltonian is characterized. Computing the self-consistent part, the real-space electron density is determined from the density matrix and from that the Hartree potential is computed - by means of FFT from the Poisson equation. The density matrix of is calculated by diagonalization of the Hamiltonian [4].

The non-self-consistet part uses two-center approximation - the matrix depends only on the distance between two atoms (their overlap between Slater orbitals for both sites), while they are not influenced by the distances of the other atoms. The non-self-consistent part of Hamiltonian is expanded in a basis of local atomic orbitals (LCAO) in spherical harmonics Y_{lm} and radial function R_{nl} [4]:

$$\psi_{nlm}(r) = R_{nl}(r)Y_{lm}(\hat{r}). \tag{1}$$

The device system gains the density matrix using NEGF, as already mentioned. This is true also for the transmission, thus for the IV-Characteristics [5].

The electron density is calculated by the occupied eigenstates of the KohnSham Hamiltonian, using the occupation f_α of the level α (Fermi-Dirac distribution) [5]:

$$n(r) = \sum_\alpha f_\alpha |\psi_\alpha(r)|^2 \tag{2}$$

This density is divided into a left and a right contribution. Using the NEFG method, left and right density matrix contributions are calculated [6]:

$$D^{left} = \int \rho^{left}(\epsilon) f(\frac{\epsilon - \mu_L}{k_B T_L}) d\epsilon \tag{3}$$

for the left contribution and accordingly for the right contribution. The Fermi function f ruled by the electron temperature T_L states for the the electron distribution in the left (right) electrode, while the non-equilibrium electron distribution governs the central region. The function $\rho(\epsilon)$ is the spectral density matrix using the retarded Green's function G and broadening function Γ and is defined as [6]

$$\rho^{left}(\epsilon) = \frac{1}{2\pi} G(\epsilon)\Gamma^{left}(\epsilon)G^\dagger(\epsilon) \tag{4}$$

for Green's function and broadening function

$$G(\epsilon) = \frac{1}{(\epsilon + i\delta_+)S - H} \tag{5}$$

$$\Gamma^{left} = \frac{1}{i}(\Sigma^{left} - (\Sigma^{left})^\dagger) \tag{6}$$

using infinitesimal positive number δ_+, Hamiltonian matrix H and overlap matrix S, both of the entire system, and the left/right electrode self energy Σ [6].

Finally, knowing the retarded Green's and broadening function for both electrodes, the transmission spectrum is computed as [6]

$$T(\epsilon) = G(\epsilon)\Gamma^{left}(\epsilon)G^\dagger(\epsilon)\Gamma^{right}(\epsilon) \tag{7}$$

2.1. Benchmark

The first step for a numerical simulation of PANI interacting with gas is a reliable construction of the PANI variation, emeraldine salt. This variation was modelled with known molecule parameters [3] and the density of states (DOS) and bandstructure were compared. After that, a device consisting of emeraldine salt was built and its transmission spectra were computed.

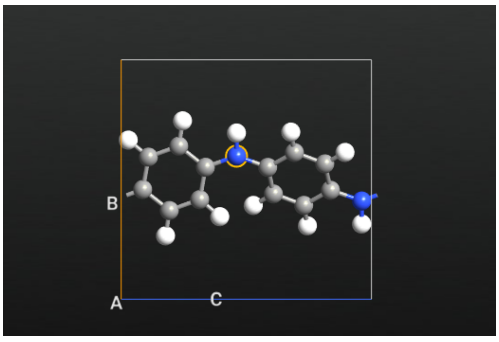


Fig. 2. One mer of emeraldine salt with optimized angles between atoms, nearing the values from [3]. The upper nitrogen atom is doped and the lower is not.

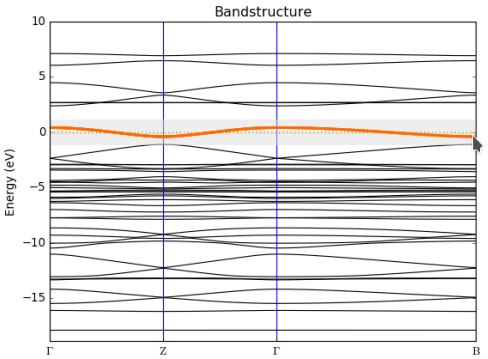


Fig. 3. Bandstructure of PANI computed as endless chain.

As the result of a chain length optimalization, a 32-mer molecule as a basic device was chosen. With this physically meaningful PANI model, IV characteristics of PANI device with and without the presence of NH_3 were studied.

3. Results

According to the emeraldine salt, every second nitrogen is doped (in this computation, every upper nitrogen is

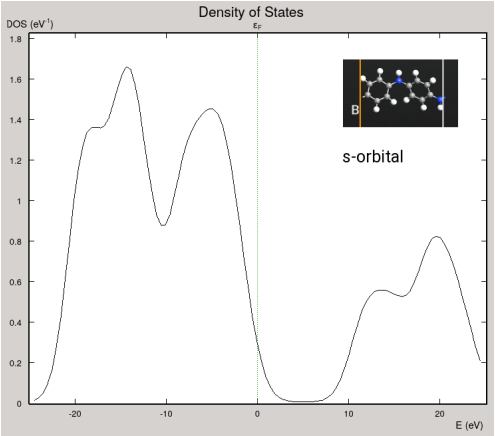


Fig. 4. DOS of a PANI endless molecule - only s-orbital.

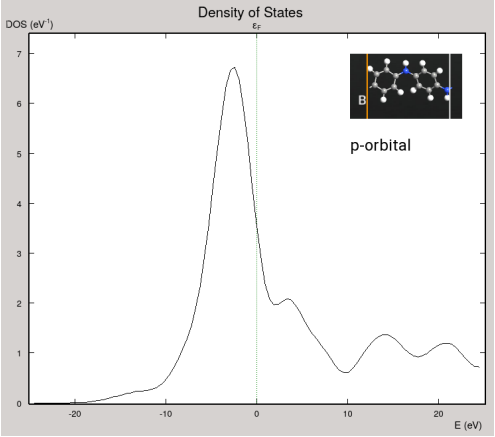


Fig. 5. DOS of a PANI endless molecule - only p-orbital, showing π-electrons contribution to conduction.

doped, see Fig.2). The interacting gas molecule can be positioned near the doped nitrogen atom, near the second, undoped nitrogen atom, near the benzene center in the middle and somewhere else. Next, the gas molecule orientation was varied. These three space positions of the molecule together with different molecule orientations were computed, and also one position chosen free from energy minimazing optimalization of position and orientation.

For all of these numerical experiments, transmission spectra and paths were computed. From these data, the IV characteristics were derived.

First a distance effect on PANI conductivity was studied. Varying the distance of the gas molecule to the chain, data as in Fig.7 were obtained. Orientation of the NH_3-molecule with its nitrogen atom facing to the doped nitrogen atom of PANI chain was chosen and the distance was measure between the center of both atoms. As seen from the figure, for distance 1.5 $\mathring{A}$ and more, no effect on the PANI conductivity is observable. Is the gas molecule nearer to the chain, a conductivity-lowering effect appears. This effect has a minimum for distances between 0.9 - 1.0 $\mathring{A}$ - the saturation current for 0.25 V reaches only 4 μA in opposition

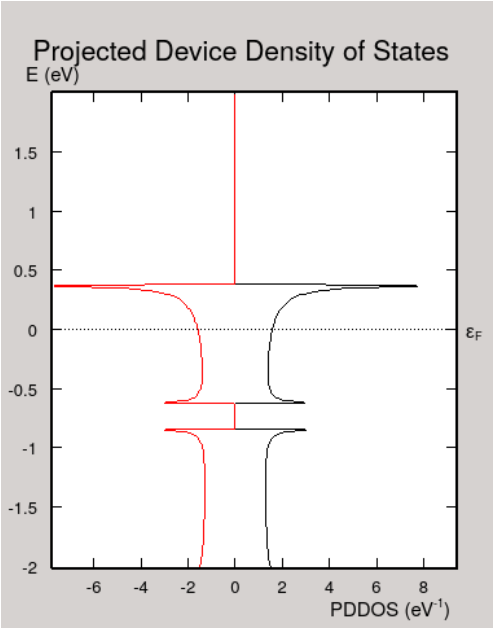


Fig. 6. Density of states (DOS) of PANI device computed on 32-mer chain, for zero bias and no gas molecules.

to PANI without any gas molecule, reaching for the same voltage 14.1 μA.

Distances under 0.9 $\mathring{A}$ are rather discutable because in this case, the nitrogen atom of NH_3 is very close to the hydrogen atom of NH-binding of the PANI chain. Thus, an optimalization of the distance is important in order to find the gas molecule position which can occure in the nature.

Next, the position of the gas molecule may vary as well as its orientation. Therefore, an molecular dynamics optimalization was chosen to find the molecule position and orientation.

As can be seen from the Fig.8, different slope and asymptotic values of current are clearly visible for various positions and orientations of ammonia molecule. Comparing the IV characteristics of PANI device without gas molecule (green color) and with optimized position and orientation of the gas molecule near the doped nitrogen atom (brown) and near the benzene center (light blue), the current value while reaching saturating voltage 0.25 V is for PANI with NH_3 more than 12% lower. Namely, PANI influenced by the optimized NH_3-molecule position near the doped nitrogen atom (center-center distance of the nitrogen atoms is $1.9\mathring{A}$) reaches a current of 12.4 μA, similar to PANI influenced by the optimized NH_3-molecule position near a benzene center (distance of the nitrogen atom to the plane of the benzene center was $4.0\mathring{A}$) reaching a current of 12.3 μA. But PANI without gas molecule reaches a current of 14.1 μA, see Tab. 1.

It is clear from this data, that such an effect on conductivity is measurable, thus a gas sensor built from PANI polymer may be possible to use.

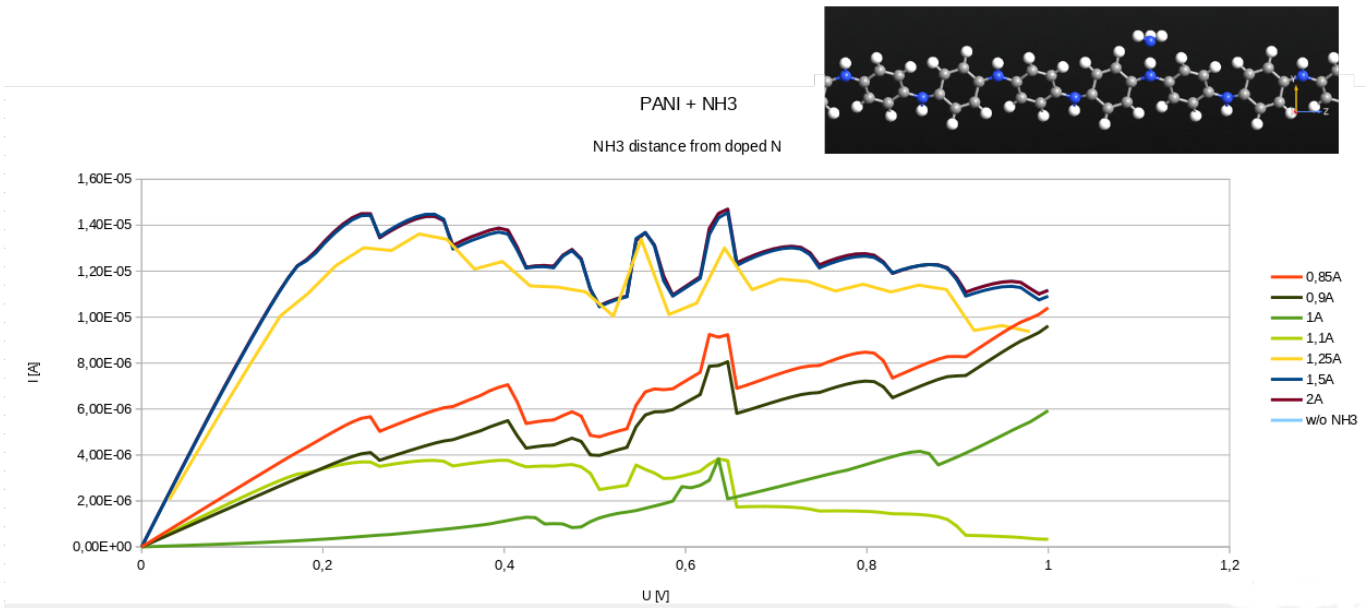


Fig.7. IV characteristics of PANI device computed on 32-mer chain.

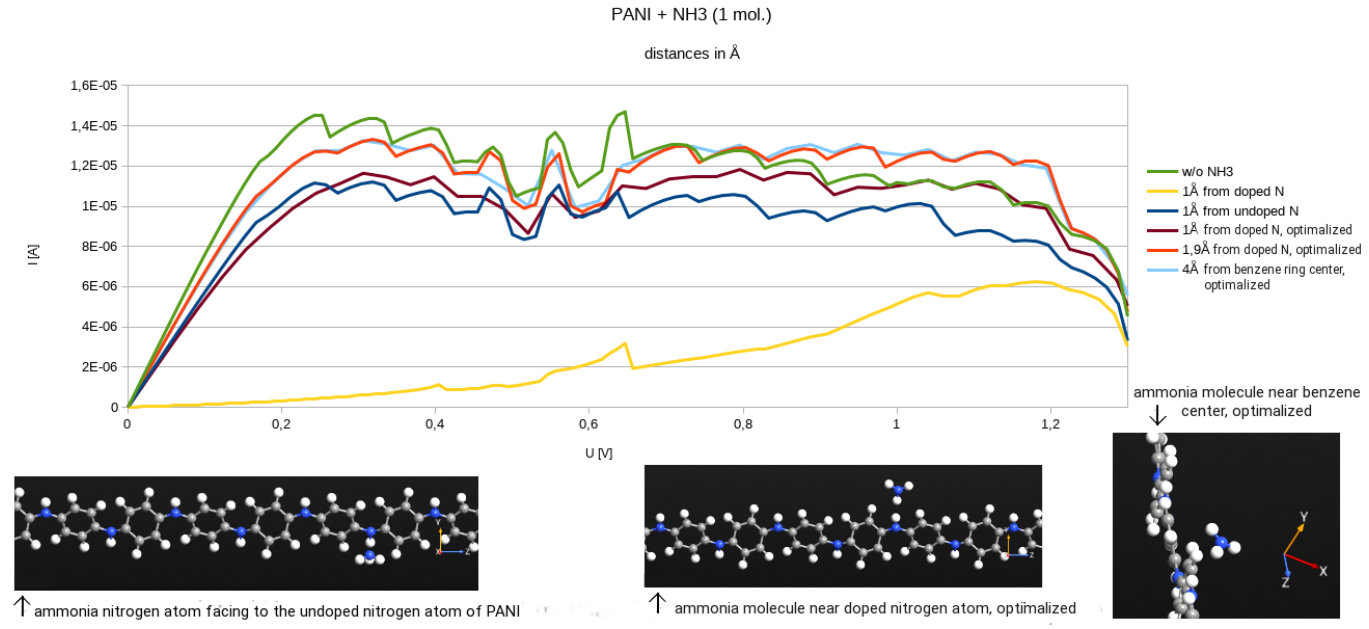


Fig.8. IV characteristics of PANI device computed on 32-mer chain.

4. Conclusions

In this paper, the electrical properties of a polyaniline chain were studied. It was verificated that a polyaniline variety, emeraldine salt, can be used as an electrical conductor - its π-electrons contribute to the conductivity. From the endless PANI chain, a device was built with electrode areas and central region, where a gas molecule of NH_3 was placed. Transmission spectra for various setups of PANI device with and without ammonia molecule in the central region of device were computed. From these data, the IV characteristics of PANI were gained. IV characteristics have shown that there is a saturation current for pure PANI device without NH_3 molecule, occuring at about 0.25 V. This current reaches over 14 μA. Next, the IV characteristics of PANI device with ammonia molecule were studied, showing that the presence of the ammonia molecule tends to lower the transmission of electrons through the device, lowering the con-

NH_3-position	voltage $[V]$	current $[\mu A]$
without NH_3	0.25	14.1
$1.9\mathring{A}$ from doped N	0.25	12.4
$4.0\mathring{A}$ from ring center	0.25	12.3
$1.0\mathring{A}$ from doped N	0.25	10.5

Tab. 1. Saturation current for various positions of ammonia gas with respect to PANI chain.

ductivity of PANI. From many possibilities, how the NH_3 molecule can be positioned and oriented, several cases were chosen, after the system of PANI chain and gas molecule was optimized for energy minimum. For these optimized states, the electrical conductivity tends up to copy the IV characteristic of pure PANI without gas, saturates at similar voltages but reaches lower currents, which are about 12.3 - 12.5 μA. This value is 12 - 13 % lower than the saturation current of pure PANI. This difference should be enough for a NH_3 sensor design as is endorsed by experimental results [7].

Acknowledgements

Research described in the paper was supervised by RNDr. Jan Voves, FEL CTU in Prague.

References

[1] MANDAL, S., SAHA, S.K., CHOWDHURRY, P. Synthesis and characterization of polyaniline Based materials: their biological relevance - An Overview. *Int.J.Curr.Microbiol.App.Sci.*,2017, vol. 6, no. 5, p. 2309-2321.

[2] GUO, Z., LIAO, N., ZHANG, M., XUE, W. Theoretical approach to evaluate graphene/PANI composite as highly selective ammonia sensor. *Applied Surface Science*, 2018, vol. 453, p. 336-340.

[3] BURGANOVA, R. Investigation of conducting polymers by computer simulations. URL: https://slideplayer.com/slide/9709152

[4] STOCKBRO, K., PETERSEN, D. E., SMIDSTRUP, S., BLOM, A., IPSEN, M., KAASBJERG, K. Semiempirical model for nanoscale device simulations. *Physical Review B Condensed Matter*, 2010, vol. 82, no. 7.

[5] QuantumATK, 2019 [cited 10.03.2019], URL: https://docs.quantumatk.com/manual/ATKSE.html

[6] BRANDBYGE, M, MOZOS, J.-L., ORDEJÓN, P., TAYLOR, J., STOCKBRO, K. Density-functional method for nonequilibrium electron transport, *Physical Review B*, 2002, vol. 65, no. 16, p.165401-165418.

[7] KROUTIL, J., LAPOSA, A., VOVES, J., DAVYDOVA, M., NÁHLÍK, J., KULHA, P. HUSÁK, M. Performance Evaluation of Low-Cost Flexible Gas Sensor Array with Nanocomposite Polyaniline Films, *IEEE Sensors Journal*, 2018, vol. 18, no. 9, p. 3759-3766.

About Authors...

Hana ŠUSTKOVÁ was born in Trutnov in 1990. She has graduated in Physics - Macromolecular physics and Geophysics - at the Charles University in Prague, now she is a PhD student of FEL, CTU in Prague, Dept. of Microelectronics.

Characterization of plasma in the PF-1000 plasma focus experiments

Pavla BEDNÁŘOVÁ, Lucia ČABROVÁ

Department of Physics, Czech Technical University, Technicka 2, 166 27 Praha, Czech Republic

bednapa3@fel.cvut.cz, cabroluc@fel.cvut.cz

Abstract. *In this paper we evaluate plasma parameters based on data obtained from experiments performed on the PF-1000 device. The PF-1000 is a plasma focus facility at the IPPLM (Institute of Plasma Physics & Laser Microfusion) in Warsaw.*
First, we explain plasma focus principles and specify parameters of the PF-1000 device. Then we describe the diagnostic tools which were used to collect the data, namely: interferometry, fast X-Ray pinhole camera, and scintillation detectors. In addition, we present a method to determine plasma density and ion temperature using interferograms. Subsequently, selected plasma parameters are introduced and calculated.

Keywords

Plasma focus, plasma parameters, plasma diagnostics

1. Plasma focus

Plasma focus is one of Z-pinch configurations. There are two main types of plasma focus devices which differ in geometry: Mather-type and Filippov-type [1]. Our experiment was performed on the Mather-type plasma focus, therefore the following text is concerned with this particular configuration. Plasma focus consists of a central electrode (usually anode), and several outer electrodes (usually cathodes). The coaxial geometry of the plasma focus is illustrated in figure 1. This system is placed inside a chamber filled with a low-pressure gas. When capacitor bank is discharged, an electrical breakdown close to the insulator occurs. The breakdown ionizes the gas between the electrodes. As a consequence, a current starts to flow through the ionized area from the central electrode to the outer electrodes forming a current shell. The current shell is accelerated by the Lorentz force which pushes the shell towards the open end of the electrodes. When the end is reached, the current shell implodes towards the axis creating a column of hot dense plasma, a so-called pinch. An increase of the density and the temperature of the plasma enables nuclear fusion reactions to occur, especially in the case of deuterium or tritium filling.

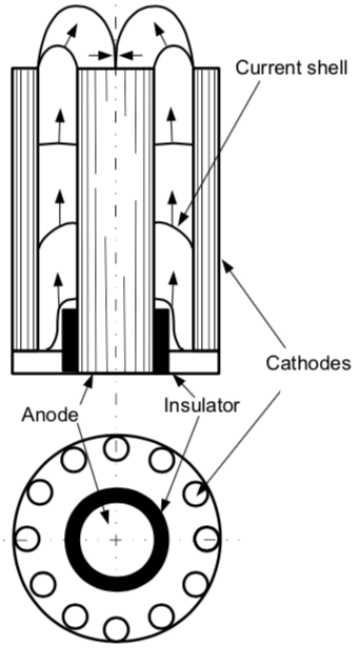


Fig. 1. Schematics of a Mather-type plasma focus device [2].

1.1. PF-1000 device

The basic parameters of the PF-1000 are summarized in the table below. The charging voltage that we used during our experiment was 16 kV which corresponds to 168 kJ of stored electrical energy in the capacitor bank.

Total capacitance	1.32 mF
Total inductance	8.9 nH
Characteristic impedance	2.6 mΩ
Charging voltage	16 - 40 kV
Stored electrical energy	168 - 1056 kJ
Maximum short curcuit current	15 MA
Maximum load current	2.5 MA
Rise time (quarter period)	5.4 μs

Tab. 1. Parameters of the PF-1000 device.

2. D-D fusion nuclear reaction

Our experiment was performed using deuterium gas at the pressure of 0.9 Torr (120 Pa) to fill the chamber. During the plasma focus discharge, the deuterium gas is turned

into deuterium plasma. The created plasma implodes by the pinch effect enabling deuterium nuclei to fuse. The D-D reaction has two branches. Both of them occur with an approximately equal likelihood. The reactions are as follows [3]:

$$ {}^{2}_{1}\mathrm{D} + {}^{2}_{1}\mathrm{D} \rightarrow {}^{3}_{2}\mathrm{He}\,(0.82\,\mathrm{MeV}) + {}^{1}_{0}\mathrm{n}\,(2.45\,\mathrm{MeV}) \quad (1) $$

$$ {}^{2}_{1}\mathrm{D} + {}^{2}_{1}\mathrm{D} \rightarrow {}^{3}_{1}\mathrm{T}\,(1.01\,\mathrm{MeV}) + {}^{1}_{1}\mathrm{p}\,(3.02\,\mathrm{MeV}) \quad (2) $$

3. Diagnostics

In this section we describe the diagnostic tools used to obtain measurable plasma parameters. We explain in detail the principles and functions of an interferometer, an MCP detector and a scintillation detector.

3.1. Interferometry

Interferometer is a device commonly used to observe density variations of a hot plasma. In our case, we use the Mach-Zehnder interferometer which contains two mirrors with a reflectivity of 50 % used as beamsplitters and two fully reflecting mirrors [4, 5]. Due to the first beamsplitter, the laser beam is divided into two beams with separate optical paths. The investigated object, in our case the hot plasma, is placed in one of the beam paths. The two paths are later merged with a set of mirrors at the second beamsplitter. The resultant beam is directed to a photo detector creating an interferogram.

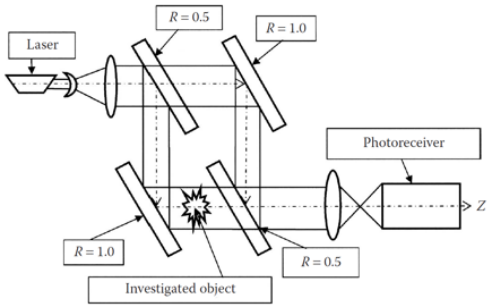


Fig. 2. Schematics of Mach-Zehnder interferometer [4].

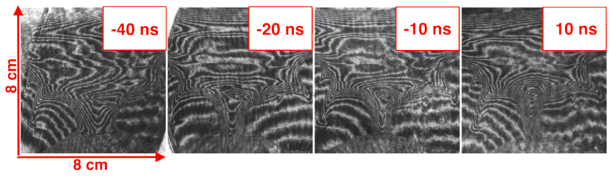


Fig. 3. Interferograms of plasma at different points of time.

Assuming a cylindrical symmetry, the following equations are used to calculate the plasma electron density n_e by analyzing the interferograms.

$$ 2\zeta \begin{pmatrix} n_i \end{pmatrix} \begin{pmatrix} 0 \\ l_{ij} \end{pmatrix} = \begin{pmatrix} \delta_i \end{pmatrix}. \quad (3) $$

Here $\zeta = -4.46 \times 10^{16}\lambda$, where λ is the wavelength of the used laser, and δ is the number of shifted strips on the interferogram. The distance which the laser beam covers in plasma l_{ij} can be expressed as:

$$ l_{ij} = \sqrt{r_{j-1}^2 - r_i^2} - \sqrt{r_j^2 - r_i^2} \quad (4) $$

where r_x stands for the distance between the axis of the symmetry and the layer with plasma electron density of n_{ex}.

As a result of the analysis described above, we obtain the dependence of plasma electron density on the distance from z-axis. Using the 16-channel interferometric system on the PF-1000 device, we obtain 16 interferograms per each shot. The exposition time is 1 ns and the time difference between two interferograms is 10-20 ns.

3.2. Fast soft X-Ray camera

The fast soft X-Ray camera is composed of a pinhole camera, gated micro-channel plate (MCP) detector, and CCD camera. The MCP detector is made as a regular array of tiny tubes or slots, densely distributed over the surface [6]. Photoelectrons are emitted by collisions of photons with a gold photocathode. These photoelectrons are accelerated in the microchannel by a high voltage. Photoelectrons cause secondary emissions of electrons which are further accelerated by electric field and also cause secondary electron emission. Thus, the electrons are multiplied in the MCP. At the end of the microchannel, electrons fly into a luminophore. Due to phosphorescence, the electron beam is converted into visible light that is captured by a CCD camera.

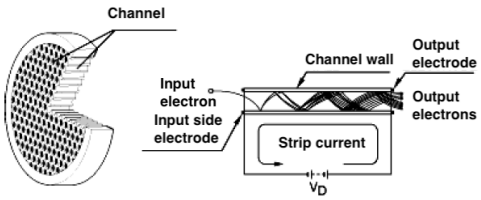


Fig. 4. Schematics of a micro-channel plate detector [7].

The photocathode is made of gold. As a consequence, the MCP is sensitive to UV and X-Ray energy region (0.1-200) nm. The MCP is supplied by high-voltage pulse which is splitted into four individual pulses that are transmitted to four independent frames of the MCP. Using different cable lengths, we can visualise the plasma in four moments. The exposure time of the MCP is about 3 ns.

3.3. Scintillation detector

When the plastic scintillator is irradiated by X-Rays or neutrons, it produces pulses of visible light. The visible light photons are detected using a photomultiplier. The photomultiplier output is coupled to an oscilloscope. Consequently,

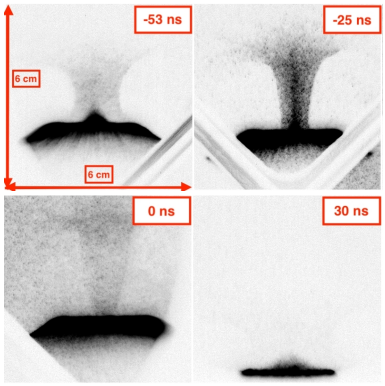


Fig. 5. MCP images at 4 different moments during discharge number 12760 .

we receive the time evolution of the neutron radiation intensity at the location of the detector. If we assume, that all the neutrons were produced at a time much shorter than the time-of-flight (ToF) of neutron from the source to the detector, we can determine neutron velocity and energy from their ToF. This method is commonly used in experiments performed on plasma foci.

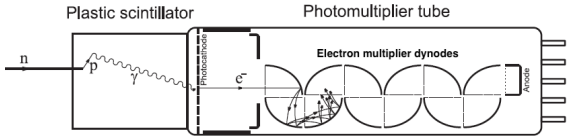


Fig. 6. Schematics of a scintillation detector [8].

Analysing the detected signal (shown in figure 7), we can determine the kinetic energy of netrons using the formula:

$$E_k = \frac{1}{2} m_n v_n^2 \tag{5}$$

where m_n stands for the mass of neutron. The velocity of neutrons v_n is calculated from the known distance between the source of neutrons and the detector and the ToF of neutrons.

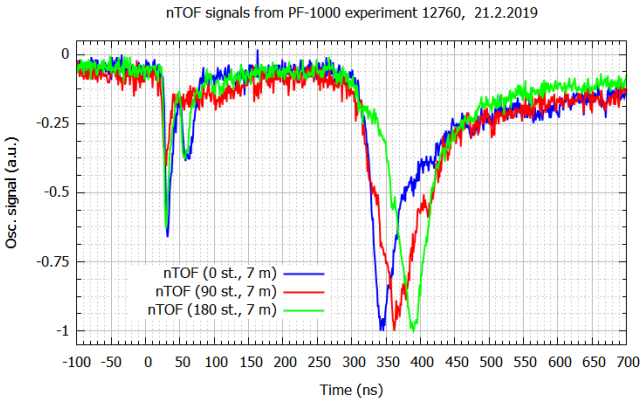


Fig. 7. X-ray and neutron signal from plasma focus discharge no. 12760 detected by 3 different scintillation detectors.

4. Plasma parameters

Plasma is by definition a quasi-neutral collection of particles containing charge carriers which exhibits collective behaviour [9]. In the following text, parameters that we use to characterize the plasma created in our experiment are introduced.

To start with, the requirement of quasi- neutrality from the plasma definition stated above can be rewritten into equation as $Zn_i = n_e$, where Z, n_i, n_e are charge state, ion density, and electron density, respectively. The atomic number of deuterium, a hydrogen isotope, is $Z = 1$.

The current I_p flowing through plasma during the plasma focus discharge creates a magnetic field. The magnitude of magnetic induction B close to the plasma column surface could be estimated by the following formula:

$$B = \frac{\mu_0 I_p}{2\pi r} \tag{6}$$

here r denotes the radius of Z-pinch column. In this experiment, we assume $r \approx 1$ cm and $I_p \approx 1$ MA.

The motion of charged particles in a magnetic field is affected by the Lorentz force. As a consequence, charged particles rotate around magnetic field lines on a diameter r_L (often referred to as Larmor radius). The frequency of the gyration is ω_c. The r_L and ω_c are given by [9, 3]:

$$r_{Le,i} = \frac{m_{e,i} v_T}{eB} \tag{7}$$

$$\omega_{ce,i} = \frac{eB}{m_{e,i}} \tag{8}$$

where the $m_{e,i}$ is the mass of electrons and ions, e is the elementary charge, and the thermal velocity v_T is given by the following expression [10]:

$$v_{Te,i} = \frac{k_B T_{e,i}}{m_{e,i}}. \tag{9}$$

Here k_B stands for the Boltzmann constant and T for the temperature. The electron temperature in our experiment is determined by spectrometric measurement as $T_e \approx 75$ keV. The ion temperature is calculated from the implosion velocity, we assume that $T_i \approx 150$ keV. The formulas for electrons and ions differ formally. Note, that the subscript "e" stands always for electrons and "i" for ions. We use this notation in this whole paper.

There are two main requirements for an ionized gas to behave like a plasma. Firstly, the Debye length must be small in comparison to the typical geometric dimension of the plasma [3, 9]. Secondly, the number of charged particles located within a Debye sphere must be far greater than one [3, 9]. To understand Debye shielding, let's assume that a positively charged sphere has been put into plasma. As a result, electrons start to flow towards the sphere creating an opposing induced electric field that shields the electric field created by the positively charged sphere. The shielding is

not ideal due to the heat energy of the electrons. The thickness of the electron layer shielding the positive potential to 1/e of its magnitude is called the Debbye length. It is given by the formula modified for deuterium plasma [9, 10]:

$$\lambda_D = \sqrt{\frac{\varepsilon_0 k_B T_e}{n_e e^2}} \tag{10}$$

Thus, the electrostatic potential arising from a microscopical non-uniformity in density is shielded within the distance λ_D. This mechanism occurs only when there are enough charged particles to create the shielding layer. The number of particles within the Debye sphere (a sphere with radius equal to λ_D) is represented by the plasma parameter N_D [9].

$$N_D = n_e \frac{4}{3}\pi\lambda_D^3 \tag{11}$$

Another parameter characterizing plasma is the plasma frequency ω_p. If there is any displacement of electrons in relation to ions in the plasma, an electric field pushing the electrons back to their neutral position is created. The electrons start to oscillate at the plasma frequency [3, 9].

$$\omega_p = \sqrt{\frac{n_e e^2}{\varepsilon_0 m_e}} \tag{12}$$

The particles in the plasma collide with each other, therefore we denote estimates for the electron and ion mean-free paths [11]:

$$\lambda_e \approx \frac{2 \times 10^{22} (T_e[\mathrm{keV}])^2}{Z n_e} \tag{13}$$

$$\lambda_i \approx \frac{2 \times 10^{22} (T_i[\mathrm{keV}])^2}{Z^3 n_e} \tag{14}$$

The plasma collisionality can also be characterized by the equilibration time between electron and ion temperature. The numerical estimate gives the following expression [11]:

$$\tau \approx \frac{A \times 10^{18} (T_e[\mathrm{keV}])^{\frac{3}{2}}}{Z n_e} \tag{15}$$

where A is the mass number of deuterium, i.e. $A \doteq 2$.

5. Results

Figure 8 shows the plasma electron density distribution in our experiment. It is evaluated at 3 different moments (10 ns, -10 ns, -20 ns). During this time period, we can see the increase of plasma electron density on the z-axis from 2.7×10^{24} m^{-3} up to to 4.2×10^{24} m^{-3}. The mean plasma density and other calculated plasma parameters are summarized in table 2.

Analysing the X-Ray and neutron signal (fig. 7) from 3 scintillation detectors positioned at 3 different angles $(0°, 90°, 180°)$ in relation to the z-axis, kinetic energy of neutrons is evaluated. The results are shown in table 3. The neutron energy differs in the 3 mentioned directions due to the directed ion beam in the high current pinched plasma. In accordance with the kinetics of the D-D nuclear reaction.

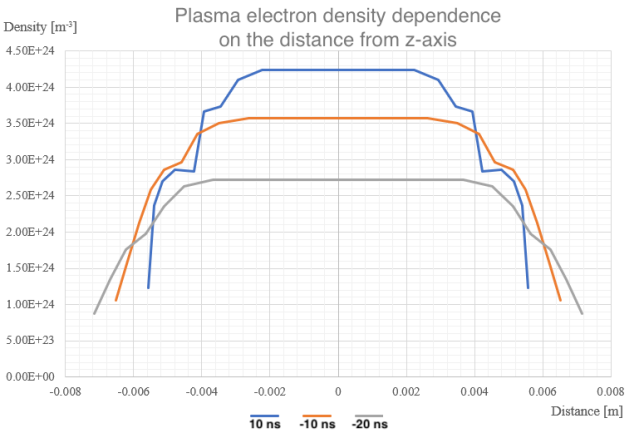


Fig. 8. Graph showing distance dependence of the plasma density for different points of time

Mean plasma electron density	n_e	2.7×10^{24} m^{-3}
Magnetic induction	B	20 T
Electron Larmor radius	r_{Le}	1×10^{-6} m
Ion Larmor radius	r_{Li}	8.9×10^{-5} m
Electron gyrofrequency	ω_{ce}	3.5×10^{12} rad/s
Ion gyrofrequency	ω_{ci}	9.6×10^{8} rad/s
Electron Debbye length	λ_{De}	3.9×10^{-8} m
Plasma parameter	N_D	682.8
Electron plasma frequency	ω_{pe}	9.2×10^{13} rad/s
Electron mean free path	λ_e	4.2×10^{-5} m
Ion mean free path	λ_i	1.7×10^{-4} m
Temperature equilibration time	τ	15.3 ns

Tab. 2. Calculated plasma parameters

Position	Neutron kinetic energy
0 deg.	2.69 MeV
90 deg.	2.45 MeV
180 deg.	2.05 MeV

Tab. 3. Neutron signal properties.

6. Conclusion

We analysed the discharge no. 12760 in the PF-1000 plasma focus experiment. The development of the discharge was captured by fast soft X-Ray camera (fig. 5). We evaluated neutron kinetic energy (tab. 3) by analysing X-Ray and neutron signal detected by scintillation detector. The neutron yield was 6.51×10^{10}. The maximum electron plasma density was determined as 4.61×10^{24} m^{-3} using interferograms. The parameters which characterize the plasma are summarized in table 2. Comparing the evaluated electron and ion mean free path with the plasma dimensions and larmor radii, we can conclude that the plasma in our experiment is strongly collisional.

These parameters are important for further theoretical considerations and applications in laboratory astrophysics. Therefore, we will continue in the investigation of the hot dense plasma in following plasma focus experiments.

Acknowledgements

The research described in the paper was supervised by prof. RNDr. Pavel Kubeš, CSc. and Ing. J. Cikhardt, PhD., Department of physics, FEE CTU in Prague and supported by the following grants: Ministry of Education, Youth, and Sports of the Czech Republic No. LTT17015 and No. CZ.02.1.01/0.0/0.0/16_019/0000778; CTU Grant No. SGS19/167/OHK3/3T/13; IAEA CRP RC-19253. The authors would also like to thank to Ewa Zielinska and Marian Paduch from IPPLM in Warsaw.

References

[1] SCHOLZ, M. *Plasma-focus and controlled nuclear fusion*. Institute of Nuclear Physics Polish Academy of Sciences, 2016.

[2] CIKHARDT, J. *High Energy Density Plasma Diagnostics Using Neutron and Gamma Detectors*. Prague, 2017. Doctoral thesis. CTU in Prague

[3] FREIDBERG, Jeffrey P. *Plasma Physics and Fusion Energy*. Cambridge university press, 2008.

[4] PERGAMENT, M. I. *Methods of Experimental Physics*. CRC Press, 2014.

[5] HARIHARAN, Parameswaran *Basics of Interferometry*. Elsevier, 2010.

[6] HLADÍK, D. *Měření neutronových zisků pomocí stříbrového aktivačního čítače*. Prague, 2018. Bachelor thesis. CTU in Prague

[7] HAMAMATSU [online]. Hamamatsu 2019. Available from: https://www.hamamatsu.com/resources/pdf/etd/MCP_TMCP0002E.pdf?fbclid=IwAR22PAmW0s9nKCIgBIvXgAS9Mg229N_RBMPL0PgcbuS4H_7TA5jCZPN8E0U

[8] ŘEZÁČ, K. *Reconstruction of Neutron Energy Spectra in Z-pinch Fusion Experiments*. Prague, 2011. Doctoral thesis. CTU in Prague

[9] CHEN, Francis F. *Introduction to Plasma Physics and Controlled Fusion*. New York: Plenum press, 1984.

[10] HUBA, Joseph D. *NRL Plasma Formulary*. Naval Research Laboratory, Washington DC Plasma Physics Div., 2013.

[11] RYUTOV, Dmitri D. Characterizing the Plasmas of Dense Z-Pinches. In *IEEE Transactions on Plasma Science*, 2015, 43.8: p. 2363-2384

About Authors...

Pavla BEDNÁŘOVÁ was born in 1997 in Benešov u Prahy, Czech Republic. She is an undergraduate student of Electrical Engineering, Power Engineering and Management at the Faculty of Electrical Engineering at the Czech Technical University in Prague.

Lucia ČABROVÁ was born in 1994 in Prešov, Slovak Republic. After completing her Bachelor's degree in Applied Electrical Engineering, FEE, CTU, she is currently a graduate student of Power Engineering at the Faculty of Electrical Engineering at the Czech Technical University in Prague.

Optimizing of experimental load of PFZ-200 plasma focus

Jan NOVOTNÝ

Department of Physics, Czech Technical University, Technická 2, 166 27 Prague, Czech republic

jenda.novotny@fel.cvut.cz

Abstract. *This paper presents a contribution to the study of neutron emission on the PFZ-200 plasma focus at the Department of physics on FEE CTU in Prague. In order to achieve the highest and the most stable neutron yields, the deuterium working gas pressure was systematically changing. We observed the plasma implosion time and discharge current by the Rogowski coil and neutron emission by the silver activation detector and scintillation time-of-flight detectors. The imploded plasma was visualized using fast x-ray pinhole camera with a gated microchannel plate detector.*

Keywords

Plasma focus, PFZ-200, neutron sources.

1. Introduction

The pulsed power discharge devices like Z-pinches and plasma foci are worldwide used in the experimental research of the plasma physics, laboratory astrophysics, and sources of neutrons, x-rays, and fast ions [1]. In the experiments presented in this paper, the plasma focus discharges with a current above 200 kA are performed in the deuterium gas. By D(d,n)^{3}He nuclear reactions of deuterons in the discharge plasma, the neutron emission with a yield on the order of 10^8 neutrons per a single shot is achieved. Similarly like in the case of other experimental devices, at the plasma foci, we cannot fully avoid certain fluctuations in the results of individual experiments. Obviously, these fluctuations are undesirable for the fundamental research. Therefore, we attempt to minimize these fluctuations and at the same time maximize the neutron yield as much is possible.

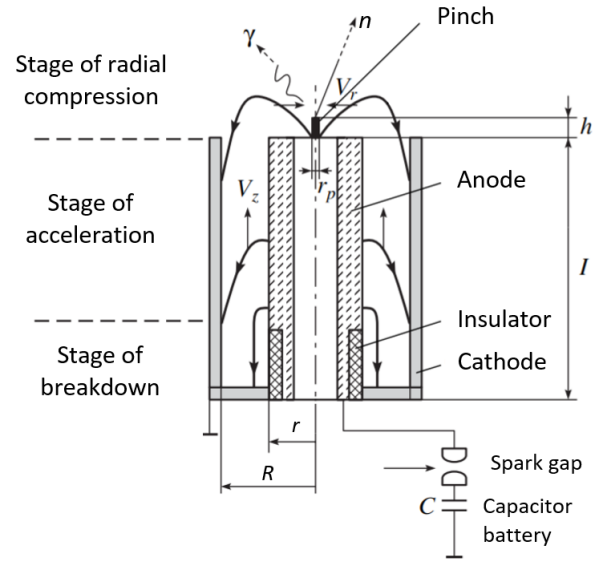


Fig. 1. Plasma focus principle scheme [2].

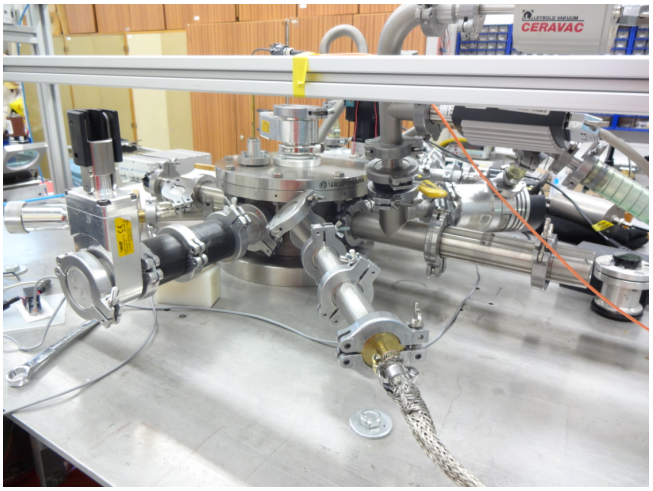


Fig. 2. Photography of the PFZ-200 device.

2. Experimental arrangement

Plasma focus PFZ-200 is a small high current discharge experimental apparatus composed of the capacitor battery, fast spark-gap switchers and anode coaxial electrode system (see fig. 2). The electrode system is placed in the vacuum chamber, which is filled by deuterium gas. The cylindrical anode is surrounded by 12 cathodes. The charging voltage of the capacitor battery is 18 kV and the maximum of discharge current reaches up to 210 kA. The current and its derivative are measured by the Rogowski coil. To evaluate the total neutron yield we use silver activation counter (SAC) [3, 4]. The SAC is located in the distance of 37 cm from the anode axis in the radial direction. In addition to this, we use scin-

tillation detectors dedicated to the neutron emission analysis by the time-of-flight (ToF) method. Whereas the ToF scintillation detectors allow us to evaluate neutron energy, for measurement of the total neutron yield the SAC is more precise. During our experiment, we use two scintillation detectors (ToF-B, ToF-C) at the distance of 269 cm and 427 cm in the radial direction from the anode axis. We are also able to visualize the discharge by the x-ray pinhole camera with a fast micro channel plate (MCP). The diagnostic arrangement is illustrated in fig. 3.

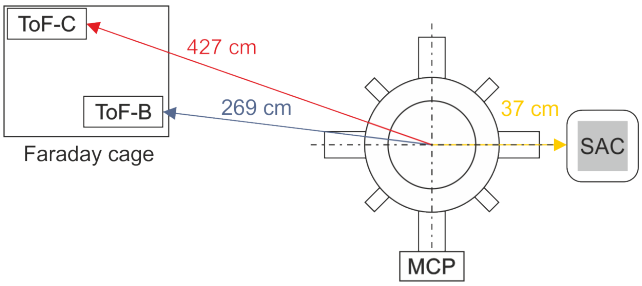


Fig. 3. Diagnostics arrangement.

3. Neutron diagnostics

In the silver activation counters, the silver foil is activated by neutrons emitted by the plasma focus. Consequently, the β^- decay radiation of the activated silver is detected by Geiger–Müller (G-M) tubes. After each shot, we count G-M pulses for one minute. The final neutron yield is calculated by formula

$$Y_n = \Phi \cdot 4\pi r^2 = (C - C_B) \cdot 4.69 \cdot 4\pi r^2, \qquad (1)$$

where Φ is a neutron flux [neutrons/cm^2], r is a radial distance [cm], C is total number of counts after one minute, C_B corresponds to G-M pulses caused by radiation background. The calibration constant 4.69 is determined by procedure presented in [5].

To evaluate neutron yield using scintillation detectors, we integrate the ToF time resolved signal. The Scintillation detector with a photomultiplier is connected to the oscilloscope. A charge Q transmitted by photomultiplier is given by

$$\mathrm{d}Q = I\mathrm{d}t, \qquad (2)$$

$$Q = \int_{t_1}^{t_2} I\mathrm{d}t = \int_{t_1}^{t_2} \frac{u(t)}{R}\mathrm{d}t, \qquad (3)$$

where $u(t)$ is a voltage measured by the oscilloscope, R = 50 Ω is the characteristic impedance of the coaxial cable and oscilloscope input. We used the known ratio between the transmitted charge of the ToF detector in the distance of 269 cm and the total neutron yield. The transmitted charge of 1 nC corresponds to the neutron yield of 8×10^6 [6]. This ratio is valid only for the photomultiplier voltage supply of 1.4 kV. For a different power supply voltage, a new calibration is required. In fig. 4, we describe signals of the ToF-B and ToF-C detectors. After the maximum of the plasma compression phase (in 10 ns) we recognise a small peak in the ToF-B detector signal. This peak represents a pulse of hard x-rays (HXR). Later, from 120 ns to 200 ns we observe a first neutron peak in the closer ToF-B detector. In the time of 270 ns, another pulse is detected. By the ToF-C detector we observe signal with a lower amplitude, because of greater distance than in the case of ToF-B. The First neutron peak comes to ToF-C later, in 180 ns. The second peak has a low-amplitude in the comparison with the signal noise.

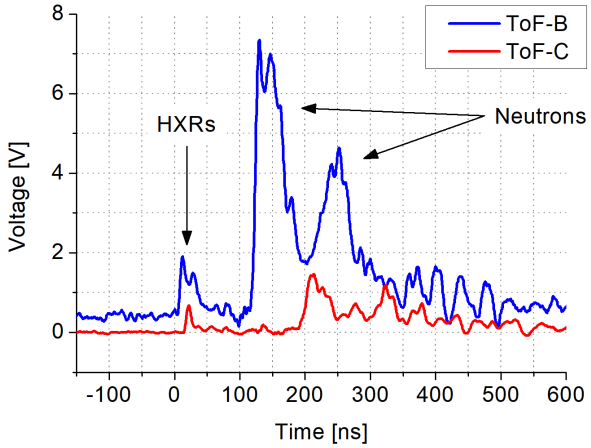


Fig. 4. Shot 18110208 - Scintillation ToF-B and ToF-C detector signals.

4. Experimental results

We performed four series of experiments with the initial pressure of 280 Pa, 300 Pa, 320 Pa and 340 Pa. For each pressure we carried out about twenty shots with the equal initial conditions. Before each shot, the chamber was evacuated up to 2×10^{-2} Pa and filled by the fresh deuterium gas, to reach the most identical initial conditions.

4.1. Neutron yield

For each initial gas pressure, the average neutron yields with their statistical variations measured by SAC and nToF are displayed in tab. 1 and 2, these values are plotted in fig. 5.

The data from the SAC and ToF-B are slightly different. The main reason is the noise in the scintillation detector signals. For example, the neutron yield measured with the ToF-B at the pressure of 300 Pa is practically doubled. Anyway, we prefer the data from the SAC and

Pressure [Pa]	Neutron yield [neutrons/shot]	Variation [neutrons/shot]	Variation [%]
280	3.4×10^7	1.8×10^7	53
300	1.2×10^8	0.4×10^8	33
320	1.5×10^8	0.4×10^8	26
340	1.2×10^8	0.5×10^8	38

Tab. 1. Average neutron yields measured by the silver activation counter (SAC)

Pressure [Pa]	Neutron yield [neutrons/shot]	Variation [neutrons/shot]	Variation [%]
300	2.5×10^8	0.9×10^8	36
320	1.7×10^8	1.7×10^8	21
340	1.3×10^8	0.5×10^8	40

Tab. 2. Average neutron yields measured by scintillation detector (ToF-B).

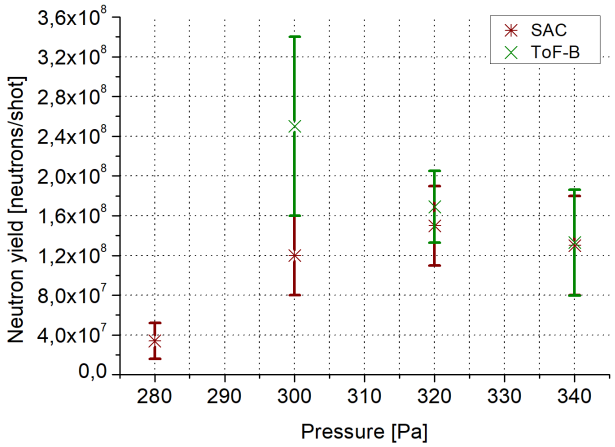


Fig. 5. Graphical view of neutron yields measured by the SAC and scintillation detector.

consider them as absolute, due to precision calibration. The scintillation detectors we use as time-resolved diagnostics, suitable for evaluation of the neutron energy by the Time-of-Flight method. Looking to fig. 5, one can identify rising neutron yield in the pressure range from 280 to 320 Pa. It seems, that at a higher pressure the neutron yield is reduced. In future, it would be appropriate to realize shots at the pressure over 340 Pa.

Based on the experimental data (tab. 1 and 2), we consider, that the highest neutron yield with the lowest fluctuation is in the shots with the deuterium pressure of 320 Pa.

4.2. Pinch dynamic

To review dynamics of the pinch plasma we measured the discharge current and its derivative by the Rogowski coil. The current maximum has a high level of reproducibility and its average value is approximately 211 kA. Only in the pressure of 300 Pa, the current maximum is about 206 kA. The maximum of the implosion phase is characterized by a steep fall of the current derivative. Thus, the plasma implosion time is given by the duration from beginning of the discharge to the minimum of the current derivative. The discharge current and the current derivative waveform are shown in fig. 6.

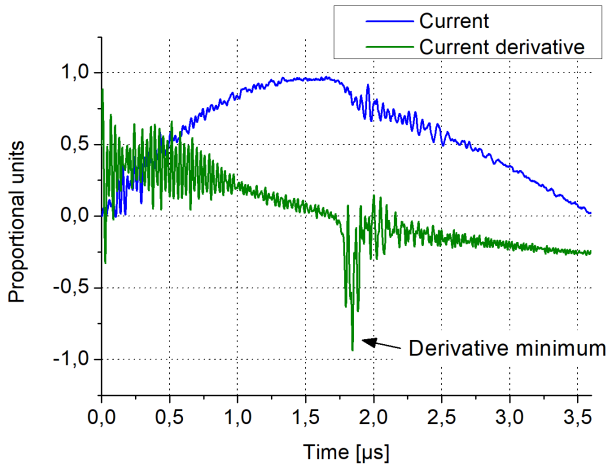


Fig. 6. The discharge current and current derivative waveform.

In tab. 3 we present the average implosion times with their statistical variations. We neglect the uncertainty of measuring instruments due to shot-to-shot fluctuation in the individual discharges. The shortest implosion times are achieved at the pressure of 300 and 320 Pa. At the pressure of 300 Pa the measured time is fairly influenced by a lower current maximum. Nevertheless, the fluctuation at the pressure of 320 Pa is halved in comparison to 300 Pa. Values from tab. 3 are also plotted in fig. 7.

Pressure [Pa]	Implosion time [μs]	Variation [μs]	Variation [%]
280	1.81	0.09	4.9
300	1.73	0.06	3.2
320	1.76	0.03	1.8
340	1.86	0.04	2.2

Tab. 3. Average time of implosion in dependence on the deuterium initial pressure.

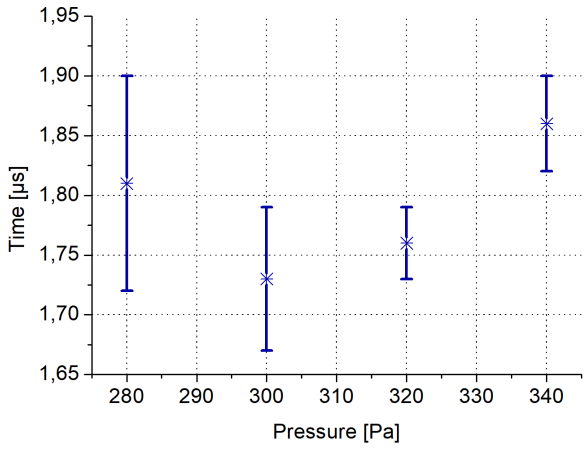


Fig. 7. Average time of implosion in dependence on the deuterium initial pressure.

4.3. Plasma visualisation

The imploded plasma is visualized by the fast x-ray pinhole camera with a gated microchannel plate detector. Fig. 8 displays the shot no. 1810208. In this figure we present four phases of the discharge with 10 ns delays between the frames. The maximum of the plasma implosion phase is observed in the fourth frame of fig. 8. There two compressions can be observed. They correspond with the two neutron pulses in the ToF signals in fig. 4.

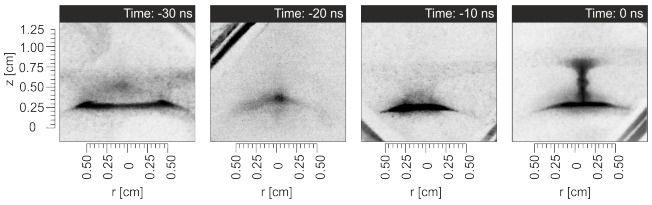


Fig. 8. Plasma visualisation by the fast MCP pinhole camera.

5. Conclusion

In the experiment on the PFZ-200 plasma focus we studied neutron emission in deuterium discharges with the current above 200 kA. The greatest neutron yield reached 2.2×10^8. Shots were performed at the deuterium gas with the initial pressures of 280 Pa, 300 Pa, 320 Pa and 340 Pa. The highest and the most stable results were obtained at the pressure of 320 Pa with the average neutron yield of $(1.5 \pm 0.3) \times 10^8$. We observed plasma implosion time dependence on the initial deuterium gas pressure. The shortest and the most stable implosion time of (1.76 ± 0.03) µs was observed at the pressure of 320 Pa. In the future, the author plans an experiment focused to neutron emission dependence on the different anode shapes at the same initial deuterium pressures.

Acknowledgements

Research described in the paper was supervised by Ing. J. Cikhardt, Ph.D.,Department of physics, FEE CTU in Prague and supported by the following grants: Ministry of Education, Youth, and sports of the Czech Republic No. LTT17015 and No. CZ.02.1.01/0.0./0.0/16_019/0000778; CTU Grant No. SGS19/167/OHK3/3T/13; IAEA CRP RC-19253.

The author would also like to thank doc. Ing. Josef Kravárik, CSc. for allowing the realization of experiments on the PFZ-200 plasma focus.

References

[1] Kubeš P.: *Impulsní silnoproudé výboje a jejich diagnostika*. Skriptum FEL ČVUT, Praha, 2004

[2] Nikulin, V.; Polukhin, S.: *Saturation of the neutron yield from mega-joule plasma focus facilities.* Plasma Physics Reports, vol. 33(4), 2007: p. 271-277

[3] Pavlát A.: *Detekce fůzních neutronů.* Bachelor thesis, ČVUT FEL, Praha 2017

[4] Hladík D.: *Měření neutronových zisků pomocí stříbrového aktivačního čítače.* Bachelor thesis, ČVUT FEL, Praha 2017

[5] HLADÍK D., PAVLÁT A., *Calibration of the Silver Activation Counter of Neutrons for Laser System and Plasma Focus Experiments*. POSTER 2017, Prague (Czech Republic), 2017, p. 1 - 5.

[6] KLIR, D., KRAVARIK, J., KUBES, P., REZAC, K., LITSEVA, E., TOMASZEWSKI, K., ... SCHOLZ, M.: *Fusion neutron detector for time-of-flight measurements in z-pinch and plasma focus experiments.* Review of Scientific Instruments, vol. 82(3), 2011

About Author...

Jan NOVOTNÝ was born in 1995 in Tábor. He is an undergraduate student of Electrical engineering, Power engineering and Management of Faculty of Electrical Engineering on the Czech Technical University in Prague since 2016. The author of this paper actively participated in three experimental internships on the PF-1000 device in the Institute of plasma physics and laser microfusion in Warsaw.

Experimental System for Classification of Defects in Vinyl Disc Records based on Typical Defects Attributes: First Proposal

Vaclav MOLDAN

Dept. of Radioelectronics, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

moldavac@fel.cvut.cz

Abstract. *This paper presents current research project focusing on description of sound attributes of different types of vinyl disc defects. Based on their typical attributes classification system is implemented and tested on samples provided by vinyl disc manufacturer GZ Media, a.s. The attributes which are described and are included in the classification algorithm are defect on each turn of the disc, defect only in right channel, echo of the defect, sequence of more clicks and leadin noise. Results 76 % successfully classified defects were achieved for defect on turn and echo of the defect, for the rest of the attributes larger test data set would be needed.*

Keywords

Audio, classification, vinyl disc, impulse noise, defects, attributes, MATLAB

1. Introduction

In recent years various studies [1, 2, 3] with topic of identification of defects and degradations on vinyl discs have been published. Also the vinyl disc manufacturers such as GZ Media, a.s. [4] also have their own control mechanism to detect defects on the vinyl discs.

The defects can be further divided into several types according to their sound characteristics and their causes. But so far no algorithm has been described which would be able to further classify and name the defect types.

Such algorithm would be useful for vinyl disc manufacturers, because it could help further investigate what was the cause of the defect and how to possibly prevent it in future.

This paper presents current research project which focuses on classification of vinyl disc records. The implemented system uses typical artefacts of each of the defects. To simplify the task, we are not dealing with the topic of defects identification, because there are already variety of solutions for this task [1, 2, 3].

In the past, classification system using KNN classification algorithm was implemented and tested for such purpose [6] but haven't received good results and therefore current research is focusing on different method.

2. Vinyl disc defects

The manufacturing process of vinyl discs consists of several phases and in each of these phases defects might be caused. These defects mainly result in clicks and pops present in the signal. But vinyl disc manufacturers can further distinguish and name the types of the defects. Each of these defects has typical sound attribute. Some of the basic defect types are described in table 1.

Name	Description
Click, pop	General defect, resulting in clicks and pops of various intensity in the audio signal. The cause might be electrostatic click, dirt in the groove, overcut, jump (skip), scratch or even error in the master records.
Scratch	There is a visible line across the vinyl disc. In case the scratch crosses the bottom of the groove, it causes click in each turn of the disc. If the scratch doesn't go deep enough to damage the bottom of the groove, it can be considered as only optical defect.
Non-fills, un-fill	This defect results in crackling in one channel (right channel). Non-fills are placed on the lee side of the edge of the matrix. In most cases, the defect crosses more grooves, therefore the crackling can be heard on each turn. The cause is the pressing, when the mass doesn't surround the matrix enough. This defect usually appears near the end of the side. We can see the non-fills as clusters of white points.
Stitching, separation damage	This type of defect is created during separation of manufacturing tools, which can result in cutting in incorrect spots on the disc.
Stitching	Other type of stitching can be created during pressing of the disc, if the disc hits the matrix.
Leadin noise, Leadin hiss	Noise is more noticeable in the leadin because of the wider grove present on leadin. It can be also caused by other factors, such as the shape of the matrix.
Humming, blowing	Extraneous sound, we can hear pulsing frequency or sound which resembles wind blowing. This artefact often occurs at the beginning of the side of the disc, it can be caused by incorrect cutting of the back side of the matrix or by incorrect form used for pressing.

Tab. 1. Examples of vinyl disc defect types ([4], modified)

3. Defects parametrization

Based on recommendation from vinyl disc manufacturer GZ Media it was decided to choose few basic types of vinyl discs defects for further parametrization, modeling and classification.

In this chapter we are focusing on description of attributes which are typical for some of the basic defect types.

3.1 Defect on turn

One of the typical artefacts is defect which occurs on each turn of the disc. This is typical for scratch of the disc, but it might occur also with other defect types (stitching, non-fills). To find such defect using algorithm we simply find the interval between consecutive defects. Because for each disc we know the number of rotations per minute (RPM), we can then calculate the interval which indicates defect on turn.

$$t=60/rpm\ [s]$$

We can't expect the interval to be always equal exactly the calculated value, because the position of the defect might not be localized exactly and also the scratch across the matrix will never be perfectly perpendicular to the grooves. Therefore, we must experimentally set tolerance level which will tell us how large deviations from the computed interval should we still accept.

3.2 Defect in right channel

Defect in right channel is typical for defect called non-fills, but it can happen with connection to other types of defects as well. The easiest way how we can recognize this type of defect is to compute RMS value of samples surrounding the sample with the defect for both channel and compare the values. In case of difference between right and left channel greater than the threshold, the defect can be considered as defect only in right channel. The threshold is set experimentally based on tests.

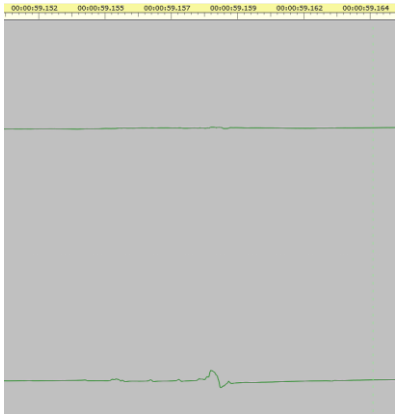


Fig. 1. Example of defect only in right channel

3.3 Echo of the defect

We can also distinguish the type of defect by the information if it's short sharp click, or if we can also hear the echo after this click.

The technique which we are using to identify such artefact is technique used generally for echo detection. Because the echo is just a delayed copy of the original signal, we can detect it using autocorrelation function.

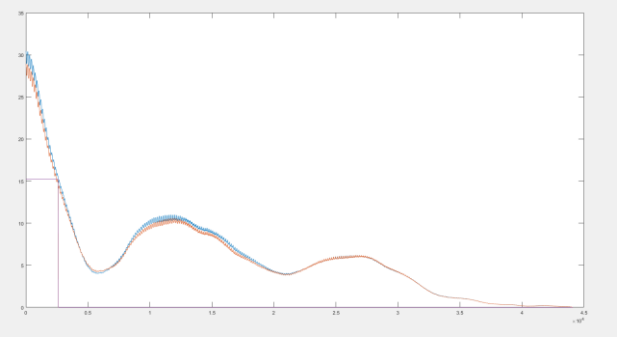


Fig. 2. Point of decay of the autocorrelation function below the threshold

3.4 Crackling, sequence of clicks

Some types of defects, mainly defect called stitching, can be characterized by not only one short click, but by crackling and sequence of more clicks positioned closely to each other. The method used for classification of such defect is to use a short part of the signal and in this segment detect clicks using click detection algorithm adopted from literature [2]. This algorithm is using wavelet transform.

After that is computed how many of the samples from selected interval contain defect and how many do not and we set ratio of these two values and compare it to experimentally set threshold.

3.5 Leadin noise, leadin humming/hiss

Leadin noise is stronger than noise present on the rest of the record because the groove is wider and therefore tends to produce more noise [4]. Also, it appears stronger because there is no other audio signal present yet. This unwanted sound can be further described as humming or noise.

The classification happens in two steps. At first the mean value of the signal in the leadin is found and is compared with certain threshold. Based on that we can say if there is any unwanted sound in the leadin. In the second step we use the autocorrelation function to say if the sound resembles periodic sound and therefore can be considered more as humming or resembles random noise.

4. Testing of the classification algorithm

For testing of the algorithm, recordings from GZ Media, a. s. were used. Together with these recordings was provided info where are the defects located and, in some cases, also what is the type of the defect.

4.1 Test data description

In total I had data about 7 recordings, and in these recordings were 122 unclassified defects, 11 examples of stitching, 22 examples of click and one example of non-fills.

In case of 3 of these recordings, information was provided whether leadin noise or huming is present or not. However, this information is provided only by graphic bar and not by numeric result.

4.2 Testing

Before classifying the type of the defect, the attributes have to be classified first (as described in chapter 3). The attributes were assigned to each defect type according to the description. The assignment of attributes to a defect type is described in table 2.

	In right channel	On turn	Sequence of clicks	Echo of the defect
Non-fills	1	1	1	1
Stitching	0	1	1	1
Scratch	0	1	0	1
Click	0	0	0	0

Tab. 2. Defect types and their sound attributes

For testing the defect present only in right channel were not enough test data available. This artefact is typical for defect of the type non-fills, and there was only one defect classified by GZ Media as non-fills present on the recordings.

When testing detection of the defect on turn different levels of tolerance of the deviation from the on turn interval were tested and good results were achieved with surprisingly high level of tolerance, 0.95 seconds, with this level we were able to get success rate 76 % of correctly classified defects.

Algorithm for detection of the echo of the defect uses three thresholds which we must set experimentally, it's the length of the segment which we work with, threshold setting the decrease level of the autocorrelation function and the length of the segment which should be above the threshold. The best results were achieved with values 1 second for length for the segment length, 0.35 for the threshold of the amplitude decrease and minimum length over the threshold as 68 %. With this combination success rate 76 % was achieved.

For testing of the leadin noise or humming there are unfortunately again not enough test data.

When testing the algorithm to detect sequence of more clicks we were unfortunately not able to achieve significant results, therefore different method would have to be used or further testing with larger data set would be needed.

On Turn	76%
Echo detection	76%
In Right Channel	Not enough test data
Sequence of clicks	Not able to successfully classify
Leadin noise, leadin humming/hiss	Not enough test data

Tab. 3. Results obtained by pilot testing of the algorithm

5. Defects modeling

As described in previous chapter, proper testing of the classification algorithm is not possible because there are not enough test data available.

Therefore, the research wants to further focus on creating models of the defects, adding them to audio recordings and then retroactively classifying them using the described algorithm.

Modeling of clicks in audio signal was already described in bachelor thesis [2]. Modeling of clicks will be further extended so that we can model all types of defects the based on the description provided in [4] and in this paper.

6. Conclusion

This paper brings proposal of algorithm to classify defects on vinyl disc records based on their typical sound attributes. Good results were achieved for attributes of defect on turn and defect with echo, for both 76 % correctly classified records, the rest of the attributes didn't provide good results yet or needs to be tested on larger data set.

The biggest issue which we are facing in this project is the lack of large test data sample. The plan for future work is to create artificial defects with audio attributes typical for each of the defect and optimize the algorithm on these idealized models of the defects. After we will be successfully able to identify models of the defects, the algorithm will be again further tested on real data.

Acknowledgements

Research described in the paper was supervised by Ing. Frantisek Rund Ph.D., FEE CTU in Prague and supported by the Grant Agency of the Czech Technical University in Prague, grant No SGS17/190/OHK3/3T/13.

Audio samples and description of vinyl disc defect types were kindly provided by GZ Media, a.s.

References

[1] BOLEK, M., *Defect detection at mechanical sound records, FEE CTU, 2016*

[2] SEMANSKÝ, M., *Detection of Impulse Noise in Gramophone Records, FEE CTU, 2017*

[3] RUND, F., VENCOVSKÝ, V., BOUŠE, J.: *Detection of Clicks in Analog Recordings Using Peripheral-Ear Model*. In: 19th International Conference on Digital Audio Effects (DAFx16). Brno. 2016, pp. 195-199. ISSN 2413-6700.

[4] List of vinyl degradations, GZmedia, internal material, 2015

[5] MOLDAN, V., *Rozpoznání titulu gramofonové desky podle krátké ukázky FEE CTU, 2017*

[6] MOLDAN, V., *Identification and classification of defects in vinyl disc records, Poster Conference, FEL ČVUT 2018*

[7] GODSILL, S, RAYNER, P., *Digital Audio Restoration. Springer-Verlag, London, 1998.*

[8] LARTILLOT, O., *MIRtoolbox 1.5 User's Manual, Finnish Centre of Excelence in Interdisciplinary Music Research, 2013*

About Author...

Vaclav MOLDAN was born in Prague, Czech Republic in 1994. In 2017 he received his bachelor's degree from Communication, Multimedia and Electronics program at Faculty of Electrical Engineering (FEE) Czech Technical University (CTU) in Prague. His bachelor's thesis was focusing on audio records identification.

He is currently studying master's program Electronics and Communication at the same faculty.

In-situ Raman spectroscopy characterization of SU-8 epoxy resin temperature dependent curing process

Alexandr Posta

Dept. of Microelectronics, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

postaale@fel.cvut.cz

Abstract. *The SU-8 is commonly used epoxy-based photoresist in manufacturing microfluidics and microelectromechanical systems (MEMS). In this paper, we investigated the temperature dependent polymerization process of SU-8 using Raman spectroscopy. The Raman spectrum were recorded in-situ during the heating of the sample from 25 to 160 degrees of Celsius. From the collected spectrum we can observe an increase in the intensity of 1183 cm^{-1} and 1108 cm^{-1} peaks corresponding to the asymmetric vibrations of C-O-C bond in polymer and decrease in the intensity of 930 cm^{-1} peaks attributed to symmetric vibration of C-O-C bond in the polymer.*

Keywords

Raman spectroscopy, SU-8, polymerization

1. Introduction

In this introduction study, we examined the changes in SU-8 photoresists by Raman spectroscopy during thermal curing. The SU-8 is a high contract epoxy-based negative photoresist with a high aspect ratio designed for application with a necessity for chemically and thermally stable films or structures. As the name suggests, it contains eight epoxy groups. The Su-8 was originally developed as a photoresist used in microelectronics for creating high-resolution masks, but recently it is used mainly in manufacturing microfluidics and microelectromechanical systems. For that application, a controlled hard-bake process is recommended to further cross-link the SU-8 structures.

The Raman effect is called after sir C.V.Raman, who first successfully demonstrated inelastic scattering of light by fluid in 1928. In Raman spectrometer, the sample is irradiated by an intense source of monochromatic light, for which he used filtered sunlight. He was awarded Nobel prize in physics in 1930 [1]. The Raman scattering can be described as an inelastic collision between an incident photon and the molecule, which result in change of molecules vibrational or rotational energy. In order for the necessity of energy conservation, the energy of the scattered photon must be different from the energy of incident photon [2]. This can be referred to as the Raman shift and is usually displayed in cm^{-1}. Various studies involving Raman characterization of SU-8 during different polymerization conditions were performed. The changes in the Raman scattering of SU-8 photoresist were studied by Suzuki et al [3].

2. Experimental

The SU-8 2 epoxy-based resin was obtained from MicroChem Inc. It was drop cast onto a glass slide with thermally evaporated 50 nm thin alumina layer for better reflexivity of the substrate. The sample was then soft-baked for 20 minutes at 60 °C at the hot plate. The Raman spectrum of the samples was collected by Renishaw inVia™ Qontor ® confocal Raman spectrometer. As the excitation source, we used 830 nm laser. After the soft-bake, the sample was placed inside the LINKAM chamber, equipped with an embedded heater, which was placed in the Raman spectrometer. The measurement setup can be seen in Figure 1.

Fig. 1. The LINKAM chamber with embedded heater inside the Raman spectrometer.

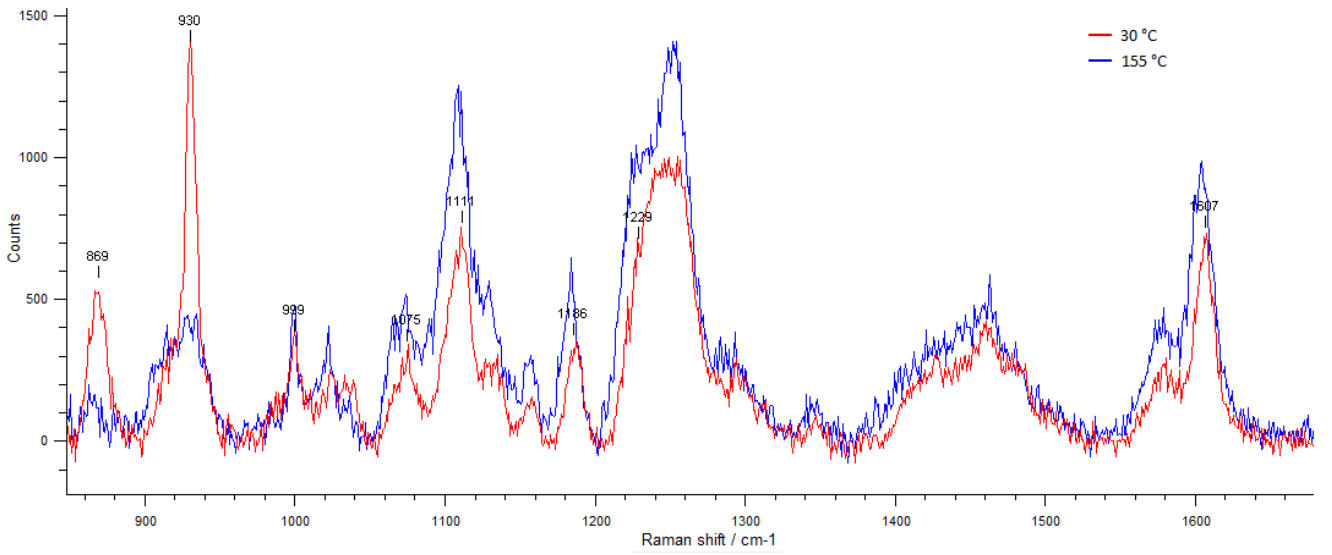


Fig. 2. The Raman spectra at the 30°C in red and in blue at the end of the hard-bake curing process at 155°C.

The Raman spectrum, we obtained in-situ during the curing process. It consisted of heat ramp from 25 to 120 °C at the rate of 5 °C per minute, then the temperature was kept at 120 °C for 20 minutes and finally ramped up to 160 °C at the same rate of 5°C per minute. The Raman spectrum was collected every minute during the heat ramp up, collecting first spectra at 30 °C. During the 120 °C period, we collected the spectrum every two minutes and every minute at the final ramp up, collecting last spectra at the temperature of 155 °C.

3. Results and discussion

In Figure 2., we can see the spectrum of the SU-8 at 30 °C in the red color and in the blue color the last spectra from the data set, collected at 155 °C. From the obtained Raman spectra we can observe the disappearance of the 869 peaks which can be attributed to the evaporation of the solvent. There is a strong decrease in the peak of 930 cm^{-1} associated to the symmetric stretching vibration of C-O-C bond and increase in the intensity of the peak around 1108 cm^{-1} and 1180 cm^{-1} which is associated asymmetric stretching vibration of C-O, C-C-O, C-O-C bonds, fracture and polymerization [4].

The 3D map of spectra obtained during the thermal curing process can be seen in Figure 3. The datasets imagine obtained spectra during the thermal curing process, with Dataset 0 corresponding to the spectra obtained at 30 °C.

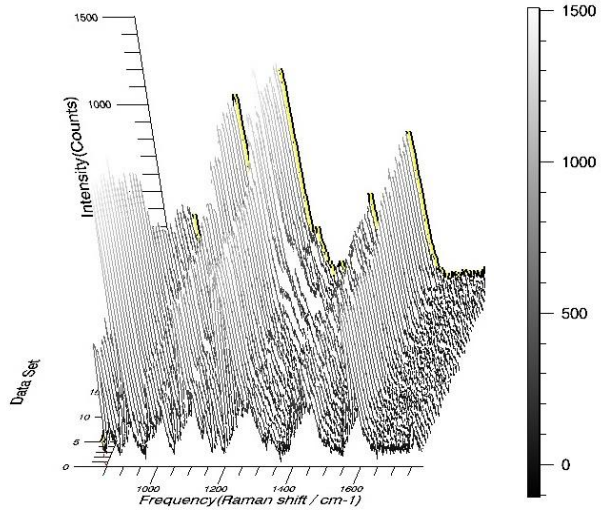


Fig. 3. The 3D map of spectra of the thermal curing process

Acknowledgments

Research in this paper was supervised by an associate professor at the Czech Technical University Faculty of Electronics, Mr. Jan Voves. This work was supported by project „Center of the Advanced Applied Natural Sciences No. CZ.02.1.01/0.0/0.0/16_019/0000778 supported by the Operation programme Research, Development and Education co-financed by European Community and by Ministry of Education Czech Republic and by the CTU student grant No. SGS17/188/OHK3/3T/13.

References

[1] Peter Vandenabeele, *PRACTICAL RAMAN SPECTROSCOPY - AN INTRODUCTION*. Jon wiley, 2013.

[2] N. B. Colthup, L. H. Daly, and S. E. Wiberley, *Introduction to Infrared and Raman Spectroscopy (Third Edition)*. 1990.

[3] D. Paipulas *et al.*, "Thermal and optical properties of sol-gel and SU-8 resists," *Adv. Fabr. Technol. Micro/Nano Opt. Photonics V*, vol. 8249, p. 82490K, 2012.

[4] Z. jian Chen, J. Yao, Q. ji Xu, and Z. hua Wang, "Two-photon polymerization fabrication and Raman spectroscopy research of SU-8 photoresist using the femtosecond laser," *Optoelectron. Lett.*, vol. 13, no. 3, pp. 210–213, 2017.

About Authors…

Alexandr POŠTA was born in Prague in 1984. In 2017, he received Master of Science degree in engineering at CTU faculty of electronic engineering and started Ph.D. study at Department of Microelectronics.

Virtual Acoustic Space Test with HMD: Pilot Test Proposal Using Unity to Create Suitable Test Structure

Viktor Jarolímek[1]

[1] Dept. of Radioelectronics, Czech Technical University in Prague, Technická 2, 166 27 Praha, Czech Republic

jarolvik@fel.cvut.cz

Abstract. *This poster shortly presents the process of design, and first testing of application, used for testing of virtual acoustic space. The paper introduces various researches that were conducted in this field and with topics related to this one. It also briefly describes the final parameters chosen for monitoring, based on previously developed tests and applications, picking the most suitable environment and approach towards the whole project. A general summary of reviewed projects and papers is also included, along with the motivation to create a tool that would allow us to compare HRTFs and obtain beneficial data that would not be a replication of previously conducted studies. The first test proposal and its structure are described in detail and a conclusion is drawn. As a first test case, it allowed us to evaluate the VR environment and get a clearer image of the next phase.*

Keywords

Virtual reality, audio quality, HRTF, Unity, HMD, virtual acoustic space.

1. Introduction

Virtual reality, in the 3D, inclusive way, is becoming more and more integrated within our lives these days and its popularity is expected to rise even more over the following years [1]. Not only there is an abundant amount of games for various platforms (Xbox, PlayStation, Head mounted displays like HTC Vive and Oculus and many more), but there are also various applications and applications for industry, learning, teaching and the list does go on. One of the key aspects for every application is sound. Among other reasons, virtual reality was built and intended for maximum immersion. In other words, people should not be able to distinguish between reality and virtual reality or at least be able to fully comprehend the environment, built in the virtual world. For that idea to fully work, all of our senses need to be included within the virtual experience. Without realizing it, all of our senses come together when we interact during every minute of our daily life. When one of these senses is off, our brain is quickly alarmed and that is when we start to question the environment, we are in. For example, that is also the reason, why we feel so comfortable, when the controllers are exactly where we see them in VR, when we reach for them. VR (in the sense of immersive VR) by itself is always based in vision, that is why we either need to use HMDs (HTC Vive, Oculus, Acer, …) or cave systems. The other important sense for us, to make it feel more real, is sound, and to be more precise, a spatial sound. For the process of learning, entertainment and a correct VR based behavior as a whole, sound that is used in these applications needs to be as precise as possible. That also includes the quality of the sound (sampling frequency, bit depth etc.), but mainly the ability for the user to localize the sound to the, preferably, exact position as one would be in real life.

The goal of this work was to create an application that would allow us and others, to some extent, measure this ability. How precise is the system rendering the location (HRTF wise) of the currently playing audio source, while also including the imperfections of human auditory system, allowing the application to have some error without disturbing the user or deviating from its purpose. Firstly, we wanted to be able to measure the quality of the sound location and its back rendering – testing general and specific HRTFs. Secondly, the goal was also to determine, how big of an error is still acceptable for the user, before it becomes disturbing to the VR experience.

2. Background research

According to [2], which in turn relies on the research of V. Larcher [3], there are, in general, three types of possible tests for measuring quality of VAS:

1. Comparison of two audio sources – the subject compares differences between two static audio sources. In my opinion, tests like these would take a long time to carry out and their results would not clearly help our case.

2. Visual representation of multiple sound sources is presented to the subject, who in turn chooses the one he/she thinks the most corresponds with the stimulus he/she is presented with

3. Absolute localization tests – the subject reports (in different ways), where he believes the audio is coming from. The reporting interface can be constructed in various

manners – anything from a simple user interface, where the subject tags the direction/area the sound is coming from, up to a 3D go-and-tag task, where the user moves around the virtual space and in order to localize the audio source.

Despite the fact that the original study was written in 2001, and the paper only a year after, my conclusion is that even to this day, the summary is fairly accurate. The same paper [2] also discusses various aspects of VR audio tests in general. One of the things, worth pointing out is that there might be a side effect (depending on the goal of the measurement) of feedback to the user. Human mind is incredibly adaptable and if we provide a feedback to the subject, after every audio source location pinpoint, a learning phenomenon might appear. Therefore, the mind itself could compensate for a mistake that was present in the test on purpose.

A more thorough division of absolute localization system is described in [4]. There are various possible reporting systems for absolute localization tests. So called "Gun in a hand pointing", when the user somehow tags the supposed audio source, while standing still, using one of the following techniques: tags using a pointing finger, using only the gaze, some external tool (eg. controllers) or using a virtual interface (eg. a sphere, a map, plains, a head model and so on). This paper conducted an experiment, based on the reporting system used in absolute localization tests. They ordered three of the methods from the least to the most effective (correct):

1. Reporting using a 2D interface on a computer screen – a three plain head model, where the subject pointed the assumed source location

2. A 3D interface on a computer screen – a head model, where the subject pointed the location

3. Pointing in a 3D environment – the sound was in a fixed location in relation to user's head, so it would move when the user turned, leaned or moved

The third option is also used in the test described later on.

Kuppanda et al. [5] created a test, where a subject was supposed to move an unknown object into their field of vision based on given audio cues. This test consisted of two possible setups. Firstly, the user did not have any visual feedback (the video feed was blacked out) and was supposed to get the audio source in the middle of his field of vision. The final position was confirmed with a mechanical click. The second setup consisted of audio cues and visual feedback, with the task of only getting the audio source inside the FOV. As a parameter, time until the button is clicked was measured for every sample. This paper, however, tested different sonification methods and therefore served as an inspiration.

A specific measurement for a minimal audible angle (MAA) was performed by Kelly and Tew [6]. They created a test with a static audio source A and a moveable audio source B. The test goal was to align both sources by moving B, until there was no distinguishable difference between them. Moreover, only the source that was currently set as active was emitting sound. The user was free to switch between them, which was done using a crossfade. The real test also didn't give any feedback, whether the B source was placed correctly or if there was a reversal error (due to the 180° confusion). However, this test was not conducted in VR and there was no visual representation apart from a 2D GUI on a computer screen. Nevertheless, this study still yielded results of the minimal audible angle between 3°and 8°, depending on sectors, laterality and longitudinality. In a concurrent experiment [7], the same group performed a test for multiple audio sources, using the other source as a distractor. Their main goal was however a bit different to mine. Fist part of the paper had to do with spectral overlapping between two audio sources. The second part used the same experiment structure as described above, while adding a distracting audio source. The task and conditions remained the same. While this study comes, idea-wise, close to my assignment, their work did not include visual stimuli and their distractor integration did not fully concur with my concept. Moreover, their final goal was to compare single-source tests and multi-source tests, where they implemented unaltered sources and sources that were modified by spectral gating, in both versions.

An entirely different study [8] created a game, where the goal of the subject was to localize a sound source as fast as possible and as accurate as possible. The environment was made up by a sphere and the sound could have been anywhere around the user. They instructed the subjects that accuracy, time and speed are all equally important and they focused on results on following questions: "How long does it take for the subject to find the source? What is the average reaction time after the sound starts playing? How fast do we rotate are head in order to find the source?" They are however not the first ones to conduct such an experiment. Fang Chen's [9] experiment consisted of localizing 3D sound, using HRTF for rendering, while focusing on accuracy and time. Moreover, he decided to separate results for male and female listeners. When combined, the average reaction time was around 5.5 seconds and the localization time was around 14.7 ± 9.8 seconds. Calle and Roginska [8] made up their test from three phases, which were basically meant to teach the subject, how the final test is going to work. The conclusion of this study mentions that this research is useful for VR content designers as it is a clear indicator of the required sound length, in order for the player to be able to safely locate it. Final results show the ability of trained subjects to locate sound source in 3.7 ± 1.8 seconds, with the average error of 15.4 degrees. But they also recommend that this is not in an average user's capabilities.

Rummukainen et. al. [10] remarked that when measuring a quality of audio in VR, it is increasingly difficult, with more degrees of freedom in movement, to separate the audio quality from quality of experience. And in the light of that, it is necessary to create new measuring methods for audio quality in VR. This study however speaks of quality of binaural reproduction systems, which

is on the brink of relevance to my task. Despite that, their test showed, structurally, the most similarities to the test that we proposed in the beginning. The user was supposed to locate the sound source, move to the assumed location and confirm the selection via controller. The audio was delivered either via open-design headphones or via loudspeakers that were set around the room. HTC Vive headset was used. This one does not allow the user to see the outside world in any way. To improve the test accuracy, the subject had to put on the headset prior entering the room in order to not be able to visually locate the loudspeakers and to actually see the extent of the experiment room. The virtual environment consisted of infinite dessert panorama and the only visual clue the subjects received was the border of the room area, depicted by a blue cage (a default Steam VR setting), where the sound source was supposed to be. After locating the current sample, a new starting location was shown and another sample was launched afterwards. The aim was to locate the audio source as fast as possible and mark the spot using HTC controller. Training session consisted of the same test, but with visual feedback of the sound locations. This whole test was conducted in 5 different specifications in regards to the delivery and rendering system – loudspeakers, binaural synthesis (also using 2 different delays), and monoaural, using a variations of intensity.

3. Test design

The main goal was to create a first environment, judge the overall feeling of the scene in VR, and try the controllers and how they react in the virtual world and asses the handling in general. Along with that, I designed a possible solution for registering the tagged locations of the perceived audio sources the user marks.

As a development tool, Unity was chosen as the most suitable. Using its engine and support for virtual reality along with extensive tutorials and forums proved to allow us to focus more on the audio quality task than on the test creation, programming wise. After setting up the whole project, the first environment was set to be a meadow, in the middle of a circle of trees and some terrain (Fig 1). To further enhance the overall feeling of the scenery, the terrain covers the visible distance so the user immersion would be high enough to ensure that the connection to the visual stimulus is high (as seen in Fig 2).

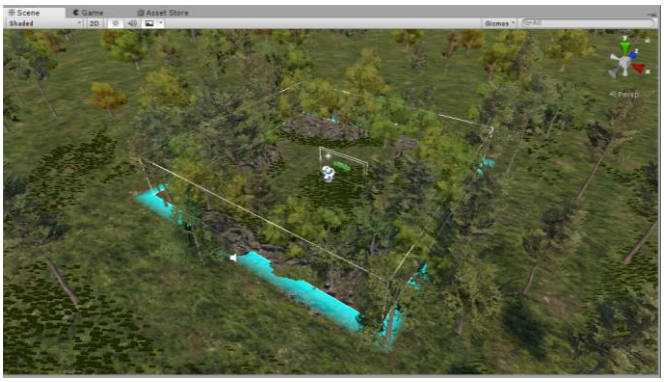


Fig. 1. A scene overview, showing the test area along with surrounding terrain.

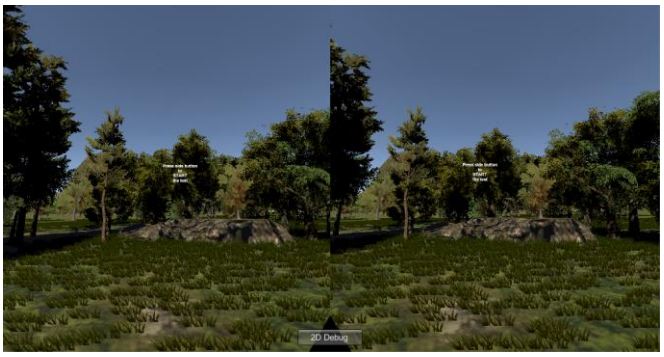


Fig. 2. Outlook on the scene design.

The audio sources were evenly distributed around the marked area – at the moment a rectangle – that was not visible to the player directly, but was marked by the more tightly grouped trees, and were played in a random order. To make calculating the marked spot of the presumed sources easier, all of them were placed directly at the border of the area as seen in Fig 3.

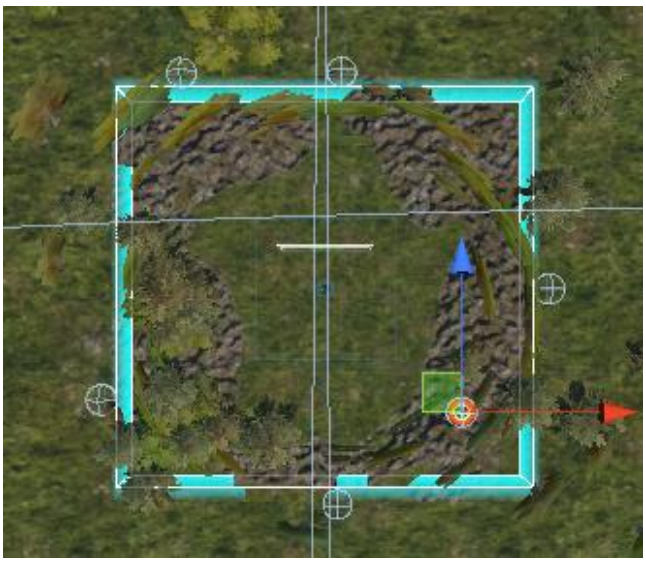


Fig. 3. Audio source distribution around the “player” area.

Marking of the audio source was realized with laser pointers, shining from both controllers in the direct line, in order to clearly see the tagged spot. A grid of colliders was constructed around the user area, to detect collisions of the rays sent from the controllers. A mechanical trigger on the controller then served as an action button to mark the perceived audio source. Fig 4 shows the realization. A different music track was picked for every source, mainly to allow repeatability by one user (me, as a developer) and also to determine, whether a song is a possible content of the audio test.

Fig. 4. Controllers with laser pointers.

3.1 Test flow

The subject is spawned in the middle of the area in Fig 3 and is asked by a prompted text to push a button to start the test. At the moment, there is no training session and all commands are given by the overseeing person beforehand. After starting the test, one of the audio sources is spawned and its coordinates are saved in a text file. The audio source keeps repeating until the user marks the perceived location via either of the controllers and this location is logged in a pair with the audio source. Afterwards a fade out is applied and a new source starts playing. This process repeats as long as there are unplayed sources and at the end a new text informs the subject that the test is over. The test then has to be shut down by the supervisor.

Multiple tests can be launched in a succession as the text file separates the individual cases. Based on the marked coordinates, number of calculations can be made – a position of the player is also known (0,0,0).

3.2 Hardware

The whole test was conducted using common middle quality headphones, as the main purpose was to test the whole environment. A better-quality headphones will be used in the upcoming tests. For the headset, various possibilities were considered, based on their availability, visual reproduction quality, resolution and overall user and research experience. As a best fit, HTC Vive HMD (not the new Pro version) was selected, as it proved to be the most reliable device for this study.

4. Discussion

Three people tested the set up in order to decide, whether a test architecture like this is viable or not. It was determined that at this moment, this test would only consist of “get the source in the field of view” and “search the location of the source in your field of view”, which might be usable, but not fully sufficient for study we wanted to achieve. Therefore, a different approach will be used in the future test, along with implementing visual distractors.

Also, a song as an audio source is not suitable for this test, because even with the spatializer, it makes the detection unnecessarily harder and disrupts the main purpose of the whole test. A more graphic content related audio will be used, to ensure the connection between the visual representation and the listening task. However, the audio still has to maintain its content value (no impulse sounds or constant-frequency tone will be used, as these are not usually heard in the real world).

Moreover, HRTF testing possibility will be implemented as well. Current test was using a generalized Unity HRTF and the following test should provide a possibility to switch for a different SOFA HRTF bank in order to compare the differences for individual test subjects.

5. Conclusion

Many studies were conducted on the topic of audio in virtual reality. Despite that, not many studies actually connect visual and audio stimuli and even less use real-life audio samples. Also, the topic of impact of visual distractors on audio has yet not been thoroughly researched and is now becoming more and more relevant as the extent of virtual reality expands. That is also the direction the next test is going to steer.

As a first test proposal, this structure was sufficient and informative, but a different architecture will have to be put in place in the final test. We were able to evaluate and compare the environment, in terms of understandability, expectations and controllability. However, a more direct approach will have to be taken with measured parameters and HRTF testing, as both of these were put aside during this test design. Moreover, Unity proved to be a fitting tool for this experiment and HTC Vive has not manifested any issues that might introduce an additional error to the test (although a clear testing area needs to be provided in order to secure safe a continuous HMD tracking).

Acknowledgements

Research described in the paper was supervised by Ing. František Rund Ph.D., FEE CTU in Prague and supported by the Grant Agency of the Czech Technical University in Prague, grant No SGS17/190/OHK3/3T/13.

References

[1] STAFF WRITER, "BusinessTech," BusinessTech, 14 10 2017. [Online]. Available: https://businesstech.co.za/news/technology/204358/the-future-of-tv-is-mobile-on-demand-and-heading-to-virtual-reality/. [Accessed 9 2 2019].

[2] J.-M. PERNAUX, M. EMERIT, D. JEROME AND N. ROZENN, "Perceptual evaluation of static binaural sound synthesis," in AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio, Espoo, Finland, 2002.

[3] V. LARCHER, in Techniques de spatialisation des sons pour la réalité virtuelle, Paris, Université de Paris VI, 2001, p. 195.

[4] J.M. PERNAUX, M. EMERIT AND N. ROZENN, "Perceptual evaluation of binaural sound synthesis: The problem of reporting localization judgements," in AES 114th Convention, Amsterdam, The Netherlands, 2003.

[5] T. KUPPANDA et. al., "Virtual reality platform for sonification evaluation," in the 21st International Conference on Auditory Display (ICAD 2015), Graz - Austria, 2015.

[6] M. C. KELLY AND A. I. TEW, "A novel method for the efficient comparison," in AES, 114th Convention, Convention Paper 5786, Amsterdam, The Netherlands, 2003.

[7] M. C. KELLY AND A. I. TEW, "The significance of spectral overlap in multiple-source localization," in AES 114th convention, Amsterdam, The Netherlands, Convention Paper 5725, 2003.

[8] J. S. CALLE AND A. ROGINSKA, "Head Rotation Data Exttraction From Virtual Reality Gameplay Using Non-Individualized HRTF," in AES 143rd convention, New York - USA, 2017.

[9] F. CHEN, "The reaction time for subjects to localize 3D sound via headphones," in AES 22nd International Conference: Virtual, Synthetic, and Entertainment Audio., Espoo - Finland, 2002.

[10] O. RUMMUKAINEN, S. J. SCHLECHT, A. PLINGE AND E. A. P. HABETS, "Evaluating binaural reproduction systems from behavioral patterns in a virtual reality - A case study with impaired binaural cues and tracking latency," in AES 143rd convention, New York - USA, 2017.

About Authors...

Viktor JAROLIMEK was born in the Czech Republic and is a master's student with major in Multimedia and Signal processing at FEE, CTU in Prague. Apart from related subject, his expertise extends to programming and testing in C#, using SQL, HTML, Typescript, React and Bootstrap.

Observer-Based Identification of Nonlinear Magnetizing Inductance of Induction Machine

Ondrej LIPCAK[1]

[1] Dept. of Electric Drives and Traction, Czech Technical University, Technicka 2, 166 27 Praha, Czech Republic

lipcaond@fel.cvut.cz

Abstract. *Conventional mathematical models of induction machine do not take into account the saturation of the main flux paths. The saturation occurs as the consequence of the nonlinear properties of the ferromagnetic circuit. Furthermore, it has been noted that the magnetizing inductance may also be strongly dependent on the magnitude of the rotor current, especially if the rotor slots are skewed or closed. This paper presents the results of a novel observer-based method of identification of the saturation of the magnetizing inductance as a function of the magnitude of the rotor flux space vector. The method is compared with the conventional single-value inductance model determined from the no-load test. The induction motor is operated in a sensorless field-oriented control where the speed is estimated by a rotor flux model reference adaptive system (RFMRAS) technique.*

Keywords

Induction machine, field-oriented control, sensorless control, main flux saturation, nonlinear modelling.

1. Introduction

Induction motors (IM) belong to the most widespread type of electrical machines [1]. This is caused mainly due to their distinct advantages such as simple construction, robustness, low maintenance, low cost and high reliability [1]. Development of modern control schemes such as direct torque (DTC) and field-oriented (FOC) control along with high computational power of today's digital signal processors (DSP) led to the massive deployment of IM in variable speed drives [1]-[2]. The modern control structures implemented within the DSPs are based on a mathematical model of the IM [1]-[2]:

$$\begin{aligned} \underline{u}_1 &= R_1\underline{i}_1 + \frac{\mathrm{d}\underline{\psi}_1}{\mathrm{d}t} + \mathrm{j}\omega_\mathrm{k}\underline{\psi}_1 \\ \underline{u}_2 = 0 &= R_2\underline{i}_2 + \frac{\mathrm{d}\underline{\psi}_2}{\mathrm{d}t} + \mathrm{j}(\omega_\mathrm{k}-\omega)\underline{\psi}_2 \\ \underline{\psi}_1 &= L_1\underline{i}_1 + L_\mathrm{m}\underline{i}_2 \\ \underline{\psi}_2 &= L_2\underline{i}_2 + L_\mathrm{m}\underline{i}_1 \\ T_\mathrm{e} &= \frac{3}{2}p_\mathrm{p}\frac{L_\mathrm{m}}{L_2}\underline{\psi}_2 \times \underline{i}_1 \end{aligned} \tag{1}$$

where the underlined symbols represent space vectors, subscripts 1 and 2 denote stator and rotor quantities, respectively, ω and ω_k denote the electrical angular speed of the rotor and general coordinate system, respectively, j denotes the imaginary unit, p_p denotes the number of motor pole-pairs, and T_e denotes the electromechanical torque. The stator and rotor leakage inductances are defined as $L_{1\sigma} = L_1 - L_\mathrm{m}$ and $L_{2\sigma} = L_2 - L_\mathrm{m}$, respectively. The magnetizing flux space vector is defined as $\underline{\psi}_\mathrm{m} = L_\mathrm{m}\underline{i}_\mathrm{m}$, where the magnetizing current $\underline{i}_\mathrm{m} = \underline{i}_1 + \underline{i}_2$. Equations (1) describe the so-called T-equivalent circuit of the induction machine which is for a squirrel cage rotor depicted in Fig. 1 and from which the definitions of the other symbols are apparent. If a constant parameter model is used within the FOC, then the magnetizing inductance L_m is considered to be non-variant. However, if saturation is included in the model, then a nonlinear characteristic describes L_m.

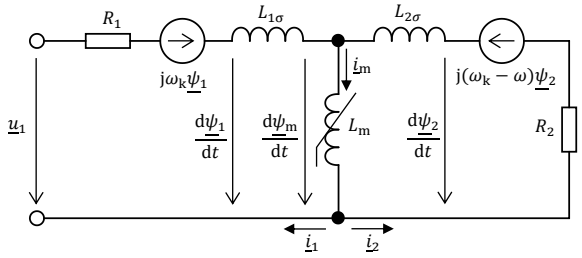


Fig. 1. Induction machine T-equivalent circuit

2. Determination and Inclusion of Nonlinear Magnetizing Inductance

Conventionally, the magnetizing inductance is measured by a simple no-load test [2]-[3]. In this way, the dependence of the magnetizing inductance with respect to the magnetizing current can be determined [2]-[3]. This paper establishes a new method of identification of the nonlinear dependence of the magnetizing inductance. The method is observer-based and utilizes equations derived from the basic IM equations given by (1). In this way, the magnetizing inductance is identified as the function of the magnitude of the rotor flux space vector, i.e.

$$L_\mathrm{m} = f\left(\left|\underline{\psi}_2\right|\right). \tag{2}$$

2.1 Measured Data

Fig. 2 shows the comparison of the magnetizing characteristic obtained from the new method and the conventional no-load test. It can be seen that both the data are in very good agreement. The advantage of the new method is that it is easy to implement and that it utilizes the same hardware that is used for the regular drive operation.

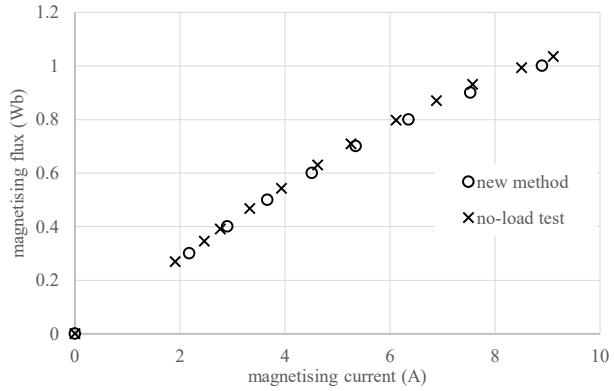


Fig. 2. Comparison of the magnetizing curves reconstructed from the new method and the standard no-load test

3. Experimental Verification

The sensorless FOC has been implemented into TMS320F28335 DSP from Texas Instruments. The clock frequency of the DSP is 150 MHz and the FOC calculation loop and the PWM period was chosen to be 100 µs. The type of FOC is the standard direct rotor-flux-oriented (RFO) where the regulation of the torque and flux producing current components is performed in the synchronous reference frame where the real axis *d* is attached to the rotor flux space vector. In such a reference system, the control of torque and flux is decoupled [1]-[2]. The drive consists of a 12-kW induction machine whose nameplate parameters along with the values used by the mathematical model measured by the standard no-load and locked-rotor tests are given in Tab. 1.

Nominal power	**12 kW**	R_1	**377 mΩ**
Nominal current	**22 A**	R_2	**225 mΩ**
Nominal voltage	**380 V**	$L_{1\sigma}$	**2.27 mH**
Nominal frequency	**50 Hz**	$L_{2\sigma}$	**2.27 mH**
Nominal power factor	**0.8**	R_{Fe}	**202.1 Ω**
Nominal speed	**1400 min^{-1}**	L_m	**82.5 mH**

Tab. 1. Nameplate and measured IM parameters

The validity of the method has been verified in the means of estimated and measured stator current. Out of (1), equations for the stator current observer in a stator-fixed coordinate system (real axis α, imaginary axis β) can be derived [2]. These equations use L_m as a parameter. Therefore, its correct value should ensure a better match between the estimated and measured components of the stator current space vector in the stator-fixed coordinate system.

Fig. 3 and Fig. 4 show the difference in terms of estimated and measured stator current for the implemented lookup-table-based compensation according to the characteristics from Fig. 2 and the conventional single-value model (Fig. 4) with the constant L_m given in Tab. 1. The reference speed was set to 100 rad·s^{-1} and the reference rotor flux value to 0.6 Wb. No external torque was applied. It can be seen that compared to the constant-parameter model the model with implemented inductance compensation significantly improves the current estimation.

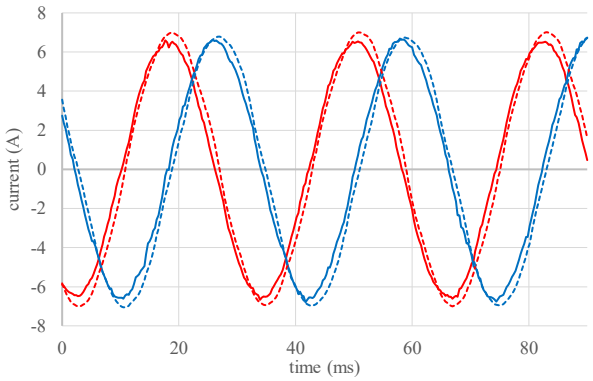


Fig. 3. Measured (solid) and estimated (dashed) components of the stator current space vector with the inductance compensation

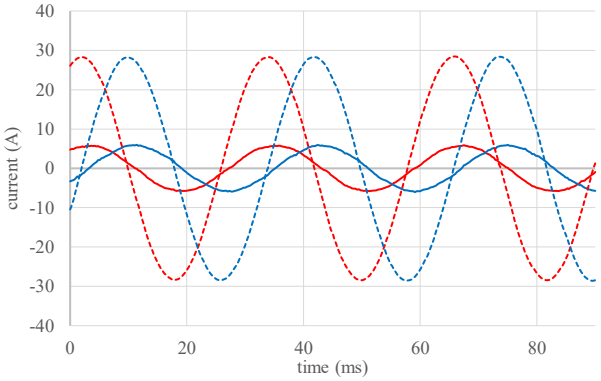


Fig. 4. Measured (solid) and estimated (dashed) components of the stator current space vector without the inductance compensation

Acknowledgements

Research described in this paper was supervised by Dr J. Bauer, FEE CTU in Prague and supported by the Student Grant Competition of the Czech Technical University in Prague under grant No. SGS19/065/OHK3/1T/13.

References

[1] PYRHÖNEN, J., HRABOVCOVA, V., SEMKEN, R. S. *Electrical Machine Drives Control: An Introduction.* 1st ed. GB: Wiley, 2016. ISBN 1119260450; 978-1119260455

[2] VAS, P. *Sensorless vector and direct torque control*. New York: Oxford University Press, 1998. ISBN 0198564651; 978-0198564652

[3] TUOVINEN, T, HINKKANEN, M., LUOMI, J. Modeling of Saturation Due to Main and Leakage Flux Interaction in Induction Machines. *IEEE Transactions on Industry Applications*. 2010, vol. 46, no. 3, s. 937-945. ISSN 0093-9994

Improvement of transient stability using STATCOM with a nonlinear PID controller

Lavr Vetoshkin

Dept. of Power Electrical Engineering, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

vetoslav@fel.cvut.cz

Abstract. *The penetration of renewables has increased over the last two decades, and consequently grid inertia has decreased. Thus modern power systems require new means of improvement of transient stability. FACTS devices (Flexible AC Transmission Systems) is a possible solution to enhance transient stability without adding costly inertia. Hence, control of FACTS has become an important topic in power system research. This paper presents an investigation of nonlinear PID controller application to STATCOM with the purpose of improvement of transient stability of a power system.*

Keywords

Transient Stability, FACTS, STATCOM, Nonlinear PID

1. Introduction

Last 20 years completely changed operation, planning and scheduling of electric grids. Deregulation and penetration of renewable energy sources complicated control of power systems. Moreover, it pushed day to day operations closer to stability limits. This transformation of electric grids also decreased spinning reserve and grid inertia [4].

Transient stability is one of the most important parameters of power systems. It can be defined as follows: an ability of generator or generators in power system to stay in synchronism with the grid during and after experiencing large disturbances. Rapid changes in parameters of the power systems, such as short circuit, connection/disconnection of power lines or power plants, are considered large disturbances. A closer look on the swing equation (1) makes clear that fundamentally for transients stability the main indicator or objective is term $\mathrm{d}\omega/\mathrm{d}t$. Hence, for a generator to stay in synchronism, the derivative term has to be equal to zero. Moreover, equations (1) and (2) explain how the electric torque is related to electromagnetic field in the stator. To stabilize the generator, only electric torque can be changed since time mechanical time constants are larger than the duration of electric transients. Clearly, equation (2) shows that electric torque can be controlled by excitation of the generator. However, sometimes it is not the best solution to rely solely on the excitation system because the system has its limits in time and amplitude of action. Thus the perception of the generator about the electric grid should be changed. Flexible AC Transmission System (FACTS) can help with that. Hence, by controlling the line voltage with a FACTS device, the electric torque of the generator can be controlled as well.

$$\frac{\mathrm{d}\omega_{mech}}{\mathrm{d}t} = \frac{1}{J}\left(\frac{P_{mech}}{\omega_{mech}} - \frac{P_{el}}{\omega_{el}}\right) \tag{1}$$

$$P_{el} = \frac{E \cdot U}{X}\sin(\delta) \tag{2}$$

where δ is load angle that is defined as follows:

$$\frac{\mathrm{d}\delta}{\mathrm{d}t} = pp \cdot \omega_{mech} - \omega_{el} \tag{3}$$

where pp is number of pole pairs.

Advances in power electronics allowed operators of electric grids to use it for improvement of control of power systems. FACTS devices are nowadays used for different purposes such as voltage control, power flow control, damp oscillations in power systems and enhance stability.

The second chapter briefly describes STATCOM (Static Synchronous Compensator) and its control system. The following chapters will outline the model and simulation results. Also, the application of variable gain PID controllers to STATCOM will be discussed.

2. Static Synchronous Compensator (STATCOM)

Static Synchronous Compensator (STATCOM) is a shunt connected Var generator, whose capacitive or inductive current output can be controlled independent of the AC system voltage, in order to control voltage or other parameters in AC power system [1]. Usually, STATCOMs utilize Voltage Source Converters (VSC), but some can be based on Current Source Converters. In the STATCOMs mostly Insulated Gate Bipolar Transistors (IGBT) and Gate Turn-off

Thyristors (GTO) are used as switching devices. Behind the VSC in the STATCOM is a capacitor or in some cases a battery. The exchange of energy between the storage device and the grid is controlled by VSC. Hence, STATCOM has an ability to inject reactive power to the system regardless of the voltage in the bus. Figure 1 displays a STATCOM with its control system. It should be mentioned that STATCOM is connected through the transformer, so the nominal voltage across semiconductor devices is reduced. Essentially, the control system of a STATCOM utilizes vector control, where the imaginary and real components of the injected current is controlled independently. The control system consists of following fundamental blocks:

- AC Voltage and Current measurements in the point of power coupling and measurement of the voltage of the capacitor on the DC side
- A phase locked loop, which provides synchronization with the grid. Also, the electric angle θ is later utilized in the Park transformation of the signals.
- AC voltage controller that utilizes Q-component of the current for voltage regulation
- DC voltage regulator controls capacitor's voltage through D-component of the current
- Current controllers that independently control D and Q components of the current. their output is voltage reference in DQ frame.
- PWM modulator which firstly does inverse Park transformation of voltage reference and then generates pulses for corresponding switches i.e transistors or thyristors

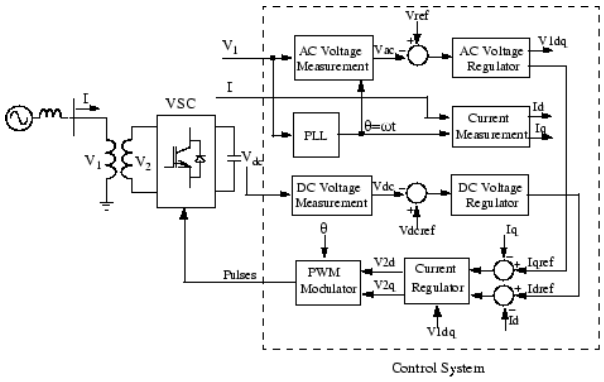


Figure 1: Single-line Diagram of a STATCOM and Its Control System Block Diagram [8]

The figure 2 displays V-I characteristics of a STATCOM. As was mentioned previously, STATCOM current injection is less dependent than the first generation of shunt FACTS devices (Static Var Compensators or SVC). Furthermore, the V-I characteristic has the slope which helps to prevent hitting the current limit frequently. For that purpose voltage droop is utilized. Hence, the voltage reference is compensated using the following formula:

$$v_{STATCOM} = v_{ref} - x_{slope} \cdot i_{STATCOM} \quad (4)$$

where x_{slope} is the droop impedance usually in the range of 0.01 to 0.05 in per unit system. Consequently, the decrease or increase in the terminal voltage for maximum possible current would be from 1% to 5% [1].

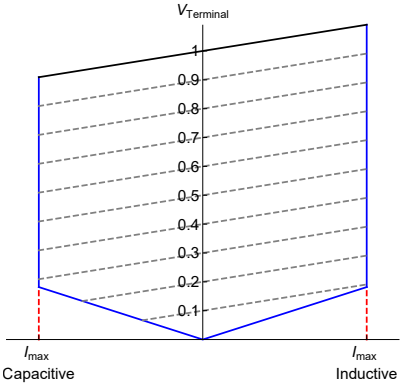


Figure 2: STATCOM V-I characteristic

3. Model

3.1. Model and Scenario

The author created the model for study of control of STATCOM during transients in MATLAB/Simulink. The model shown in the figure 3 consists of:

- A 1000 MVA synchronous generator
- A 1000 MVA 13.8kV/500kV power transformer
- Transmission lines with nominal voltage 500 kV (parameters are shown in table 1)
- Electric grid with 2500 MVA short circuit power
- A 100 MVA STATCOM placed in the middle between the grid and the generator
- A few loads: 400 MW+100 MVar, 200 MW+50 MVar, 300 MW

The scenario of the simulation is as follows:

- Before the simulation the power flow study is performed, thus the grid in simulation time 0. s is in steady state.
- At simulation time 0.1 s the tree phase fault occurs
- At 0.2 s simulation time the faulted section is disconnected

It should be mentioned that for this situation the generator will lose synchronism without STATCOM.

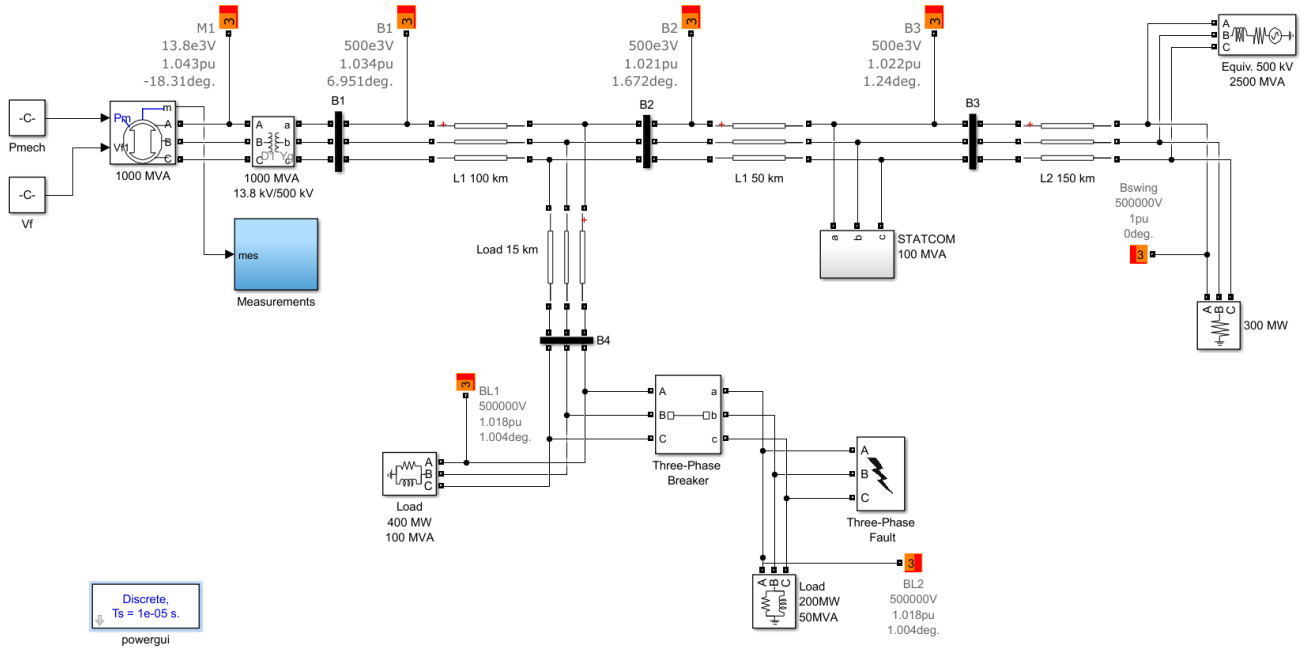


Figure 3: Simulink model

3.2. Detailed model of STATCOM with variable gain PI controller

The Author using Simulink toolbox Simscape PowerSystems implemented the detailed model of a 100 MW 48 pulses GTO based STATCOM. The line diagram of the implemented topology is shown in the figure 4. The benefits of the topology are described in depth in [1]. The zigzag connection of transformers reduces higher harmonics, by shifting phases 15 degrees. The author chose phase shift ± 7.5 degrees, hence the total shift will be 15 degrees. It should be mentioned since the transformers create phase shift the pulses should be shifted as well.

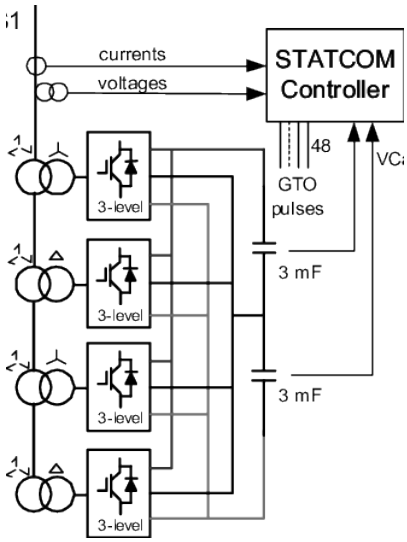


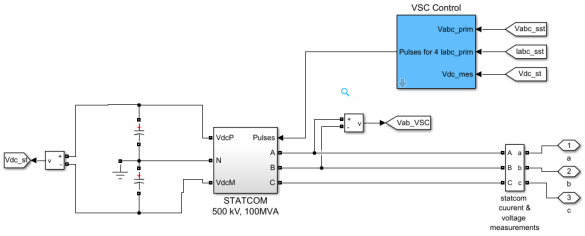
Figure 4: GTO 48 pulses STATCOM electrical scheme

Figure 5 shows the Simulink model of the 48 pulses GTO STATCOM.

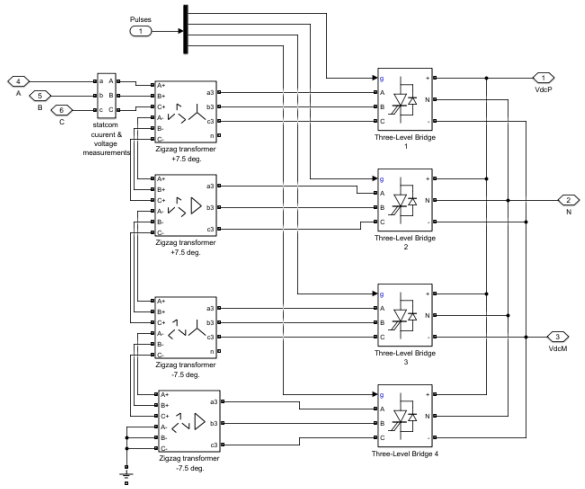
The author decided to improve the existing control system by applying a variable gain PID controller instead of classic linear PID. Moreover, the author observed that traditional PID regulator performs better during transient when it has a higher proportional gain. However, it is understandable that during high error values the integral component slows down the regulation and then after reaching the reference it still can cause overshoot. On the contrary, the proportional component of a controller has not any delay in action but cannot work without error. Hence, the author assumes that during large error the proportional gain should be higher and the integral should be suppressed, thus for small error values, it would be vice versa. Furthermore, usually the linear dependence of gains on error is used (for example in [9]), but the author chose to use nonlinear. Figure 6 shows the dependence of gains on absolute error value. It should be noted that normally in vector control of STATCOM and more generally in control systems of FACTS derivative component of the controller is not used.

3.3. Simulation results

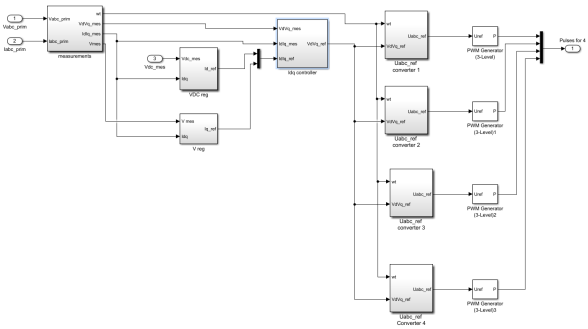
To compare the classic PI regulator and nonlinear variable gain PI, the author used the same model and tuned the regulators. The results displayed in the figures 7 and 8 clearly show that the PI controller with variable gains can be faster in restoring voltage during transient conditions. Also, there will be minor improvements in maximum load angle and slightly better damping of oscillations.



(a) STATCOM connection to capacitors and the grid



(b) Transformers and VSC converters



(c) Model of STATCOM: Control system

Figure 5: Model of GTO based STATCOM with four 3-level-VSC

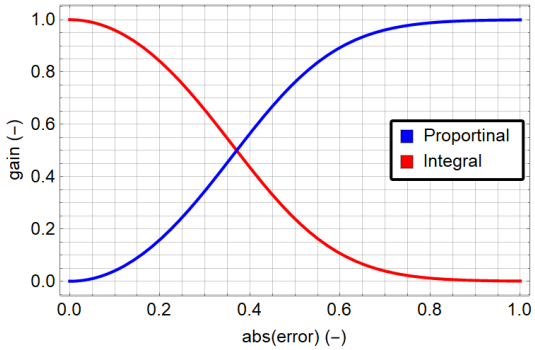


Figure 6: Variable gains

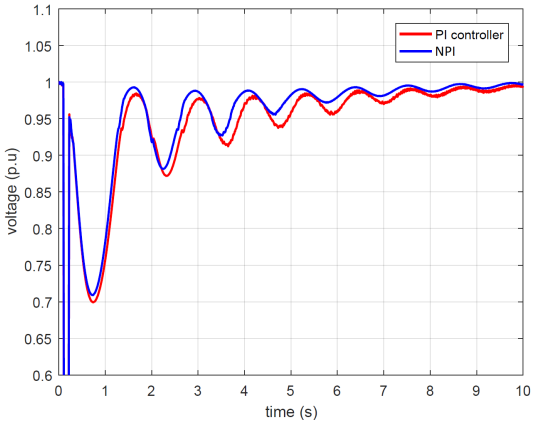


Figure 7: Voltage profile during transient

Note: NPI - Nonlinear PI controller

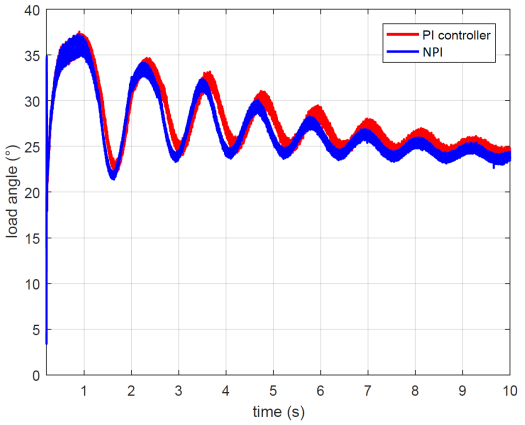


Figure 8: Load angle during transient

Note: NPI - Nonlinear PI controller

4. Conclusion

The results of simulations made by the author show that in the case of the variable gain PI controller there is a minor improvement. However, it is early to claim that it will be the case for every possible transient and system where STATCOM could be deployed. Hence, the future work will be focused first of all on the investigation of variable gain PI regulators for different systems, and transients. Secondly, the author will apply an optimization technique for tuning the regulators. Furthermore, different optimization algorithms will be considered. Thirdly, the author hopefully will research application to other FACTS devices. The author considers the work presented in this paper as a proof of concept that must be fully investigated.

Acknowledgment

Financial financial support of Czech Technical University in Prague through grant number by SGS17/181/OHK3/3T/13, is highly acknowledged.

References

[1] Hingorani, Narain & Gyugyi, Laszlo. (2000). Understanding FACTS: concepts and technology flexible AC transmission systems. IEEE Press, New York. Understanding FACTS: Concepts and Technology of Flexible AC Transmission Systems.

[2] N. G. Hingorani, "FACTS Technology - State of the Art, Current Challenges and the Future Prospects," 2007 IEEE Power Engineering Society General Meeting, Tampa, FL, 2007, pp. 1-4.

[3] K. Padiyar, FACTS controllers in power transmission and distribution, 1st ed. Tunbridge Wells: Anshan, 2009.

[4] A. Ulbig, T. S. Borsche, and G. Andersson, "Impact of Low Rotational Inertia on Power System Stability and Operation", no. Journal Article, 2013.

[5] E. Spahic, D. Varma, G. Beck, G. Kuhn and V. Hild, "Impact of reduced system inertia on stable power system operation and an overview of possible solutions," 2016 IEEE Power and Energy Society General Meeting (PESGM), Boston, MA, 2016

[6] D. V. Tien, P. Hawliczek, R. Gono and Z. Leonowicz, "Analysis and modeling of STATCOM for regulate the voltage in power systems," 2017 18th International Scientific Conference on Electric Power Engineering (EPE), Kouty nad Desnou, 2017, pp. 1-4.

[7] "Implement phasor model of three-phase static var compensator - Simulink", Mathworks.com [Online]. Available: `https://www.mathworks.com/help/physmod/sps/powersys/ref/staticvarcompensatorphasortype.html`. [Accessed: 13- Dec- 2018].

[8] "Implement phasor model of three-phase static synchronous compensator - Simulink", Mathworks.com [Online]. Available: `https://www.mathworks.com/help/physmod/sps/powersys/ref/staticsynchronouscompensatorphasortype.html`. [Accessed: 14- Dec- 2018].

[9] Korkmaz, Mehmet & Aydoğdu, Omer. (2011). A Simple Approach to Design of Variable Parameter Nonlinear PID Controller.

Apendix

Table 1: Transmission Line Parameters

	Positive Sequence	Zero Sequence
Resistance (Ohm/km)	0.01755	0.2758
Inductance (mH/km)	0.8737	3.220
Capacitance (nF/km)	13.33	8.297

About Author

Lavr Vetoshkin was born in 1995 in Tyumen, Russia. In 2013 he enrolled at the faculty of electrical engineering CTU. In 2016 he graduated with a bachelor degree in Applied Electrotechnics. He continued his studies at the master program and finally graduated in 2018 with a master degree in Power Systems. Currently, Lavr is a PhD student at the Department of Power Electrical Engineering at FEE CTU.

A new approach to investigation of Series Arc behaviour

Lavr Vetoshkin

Dept. of Power Electrical Engineering, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

vetoslav@fel.cvut.cz

Abstract. *A series arc fault is the most frequent electrical cause of fires. Currently, many in the fire safety field are concerned with deployment, developing and improvement of Arc Fault Circuit Interrupters. This paper investigates series arc characteristics in the presence of household loads. The paper presents different approaches to the investigation of a series arc's features. The standard deviation (SD) analysis of high frequency component of a current showed promising results, which may be a good indicator for the detection of a series arc.*

Keywords

Series Arc, Arc Fault, Fourier Transform, Standard Deviation

1. Introduction

According to [1] arcing faults cause more than 30000 home fires a year in the USA and many times more worldwide. Peoples safety is always a major concern; thus the fire statistics triggered the research of arc fault circuit interrupters (AFCI) decades ago. However, due to nature of arc faults the existing devices are not 100 percent reliable. Hence, the investigation into series arc feature continues. The ultimate goal indeed is to find a reliable indicator for fault detection.

Series arc can be interpreted as an unintentional discharge on the suface of a conductor. Series arc fault occurs in the path of a circuit in series with a load. Mostly, an interruption of insulation or flawed mechanical contact causes this type of fault. Series arc fault represents an extra load, in a series with existing, in the circuit. Consequently, the current, flowing into the faulted branch, is close to standard operating current. Therefore, conventional protection devices cannot detect series arc faults.

The arc itself is a very complicated phenomenon. Therefore this section introduces series arc and describes its commonly known features. AC arc was studied a lot, and its features are well known. Major influential factors are voltage, gap distance, electrode surface and line impedance [3]. The voltage applied across electrodes controls arcing, so when AC voltage crosses zero, the arc extinguishes. After restoring a critical value of voltage, arc reignites. Figure 1 shows the arc voltage waveform. Between those two processes is a current gap, which duration depends mostly on the distance between electrodes and state of the plasma. Longer the distance gap, a higher voltage is needed to initiate arcing. Plasma between electrodes affects the insulating properties of the fluid between electrodes, and it leads to a smaller voltage that can cause a breakdown. Arc extinguishing and re-ignition cause a current edge. Current raise ration depends on applied voltage needed for a breakdown, on the load, and on the circuit parameters. Figure below 2 shows these edges that are often called shoulders.

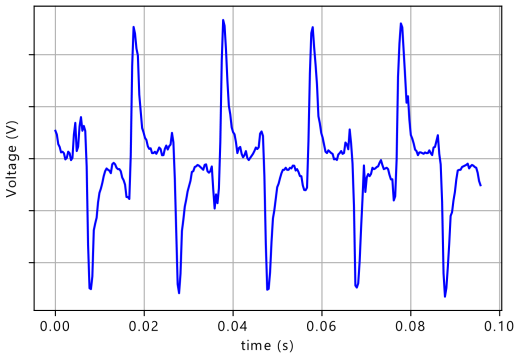


Figure 1: Arc voltage

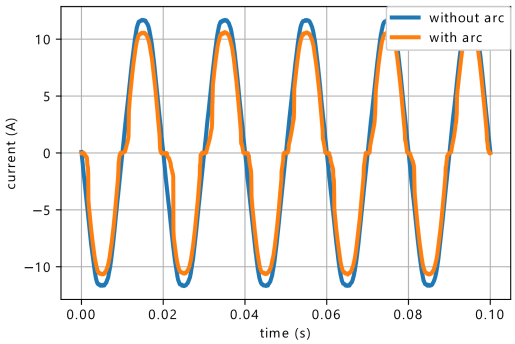


Figure 2: Current waveform without/with series arc

It is obvious that current gaps decrease RMS value of the signal. Also, time of arc re-ignition varies due to stochastic nature of the phenomenon.

Several papers outline the importance of higher harmonics assessment in arc fault detection algorithm for reliable disclosure of failure [3, 4]. Re-ignition of current, coming in hand with a current edge, generates high-frequency noise. As a result, signals spectrum significantly changes. There are different opinions which part of the spectrum suits better for detection algorithm. Carlos E. Restrepo in his paper suggests using RF component for AFDD [4]. Other researchers focus on frequencies in the range of 2 to 5 kHz [3]. Nevertheless, the idea behind those approaches is noise generated during arcing, which elevates higher harmonics.

The change in spectrum causes the decrease in amplitude of fundamental harmonic. Figure 3 shows current signals from figure 2 in frequency domain, obtained by using Fast Fourier Transform (FFT). High-frequency component is displayed on figure 4. The high-frequency component of the signal with arcing vastly distinguishes from the normal condition. Arc re-ignition, apparently, causes the peaks.

The characteristics outline features of arc presence in the circuit. The elevation of third harmonics (150 Hz) amplitude is noticeable, overall spectrum elevation up to 2.5 kHz is sharp and increase in higher frequency component is perceptible too. All that phenomena could be good and reliable indicators of an arc fault in the circuit. Unfortunately, harmonics of those frequencies are produced by different loads, for example, converters produce 3, 5 and 7 with high amplitude, there is the difference between them. First of all, the arc is a stochastic process, on the contrary power electronic devices are not. Therefore, it should be taken into account, during the development of AFDD. Otherwise, that kind of loads could cause undesired tripping of protection device.

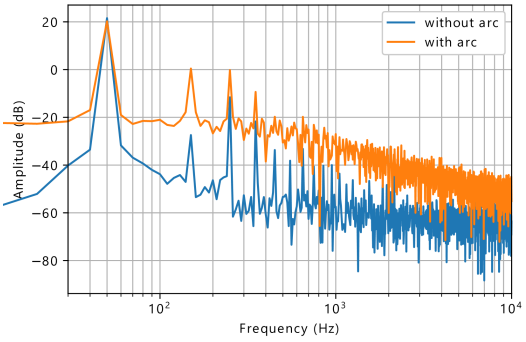


Figure 3: Spectra of current signals

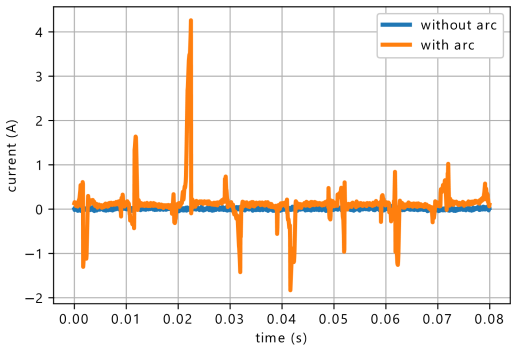


Figure 4: High-frequency components (higher than 50th harmonic) of currents

is shown in the figure 5. It explains from a high-level perspective what the setup can provide. The first block from the left is a transformer which converts the current signal into the voltage signal, which is gained after that and goes to the oscilloscope. That is a raw current signal. This signal additionally goes through two analogue filters. The first one is a second order low-pass filter, which output is gained and connected to the oscilloscope. The second is a second order highpass filter, which output is also gained, after that goes through an envelope detector and eventually to the oscilloscope input. This signal shows the envelope of the high frequency component of the current in the circuit. Furthermore, the arc's voltage is measured. The output of a transformer gained and connected to the input of the oscilloscope.

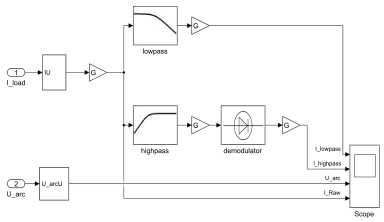


Figure 5: The block scheme of the setup

2. Measurement setup

The setup made at the faculty of electrical engineering at CTU for arc faults measurements uses arc generator. It is one of the two options how to create arcing in a circuit according to IEC standard [7]. The block scheme of the setup

Figure 6 shows the detailed setup scheme. The subscribed arrows are oscilloscope chanels. Transistor amplifiers are used to gain the signals. Two arrows mean an arc generator, which can be short circuit with the switch connected in parallel.

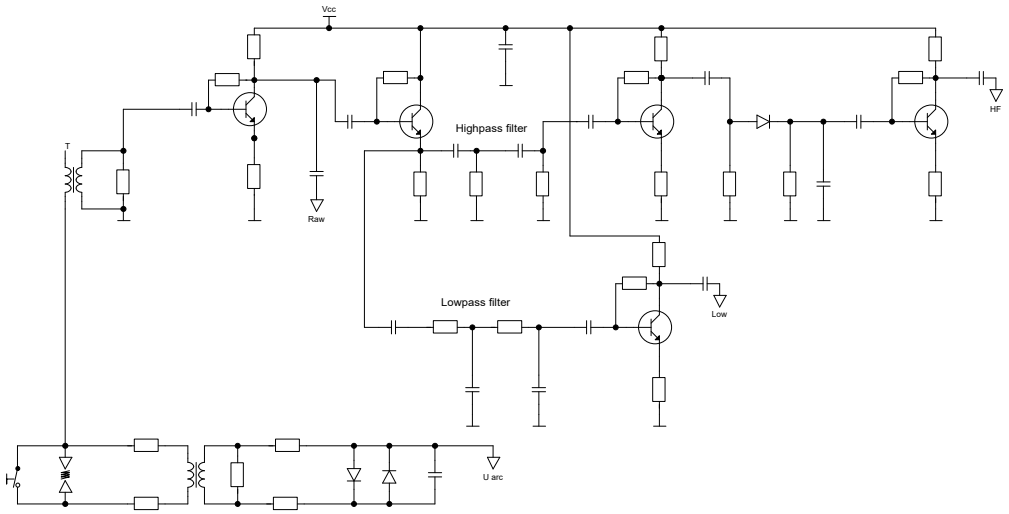


Figure 6: Electrical scheme of the setup

3. Investigation into series arc features

3.1. Phase angle

The series arcing is a stochastic process during which many electrical parameters of an arc vary. It implies that arc resistance changes in the time domain. This variation causes a change in a phase angle of a current signal in a circuit with arcing.Hypothetically, in the circuit with a series arc resistance constantly changes, hence the phase angle as well. It is similar to constant switching between every possible phase angle. The following text will show that it may be the case. However, the R/L ration in a circuit can change during normal operating conditions for example change of the temperature of a load can cause variation in electrical parameters.

By using Fourier transform, from acquired data for the circuit with the power drill as a load phase angles were extracted. The figure 7 shows calculated phase angles for different cycles of the signals. It is clear that in case of arcing phase angle of the current had been changing from cycle to cycle. However, the phase angle in the case without the presence of a series arc had been changing too. It could have happened because of numerical errors or change in load's electrical parameters. However, a variation of a phase angle does not affirm that in the circuit is a series arc; it only says that there is a change in R/L ratio. Obviously, it does not specify the cause of the variation. Phase angles of the current obtained from the circuit with a power drill without arcing, are displayed on the figure 8. In the of this record, the power drill starts, and it causes variation in the phase angle. It proves that not only an arc can cause a shift of a phase angle. Essentially, the accuracy of such calculation depends on the precision of Fourier coefficients estimation. The period for computation of that coefficients, i.e. fundamental frequency is crucial. Therefore the author tried to es-

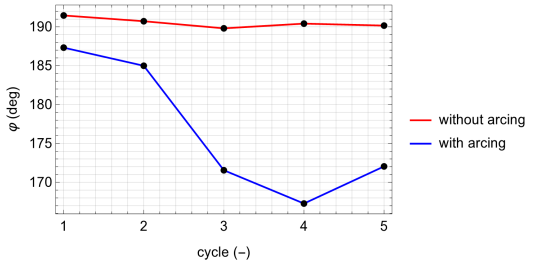


Figure 7: Phase angles of current of the circuit with power drill with/without arcing

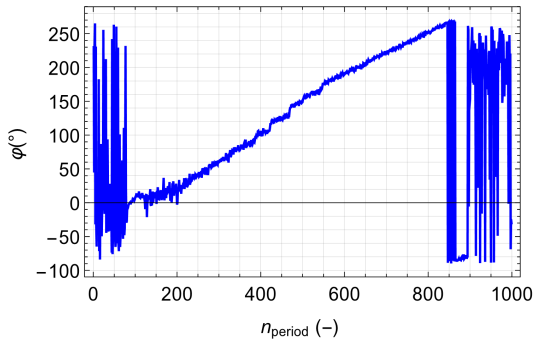


Figure 8: Phase angles of the current with a power drill without arcing

tablish what imprecision of the fundamental frequency value can be tolerated. For the same data as in the previous section, the changes of the phase angle for different fundamental frequencies were calculated. The fundamental frequency was computed by using the statistical method described in [9]. Figures 9 and 10 prove that calculation of a fundamental frequency is necessary for correct estimation of a phase angle. Also, they show that even a small deviation from true value causes a significant error.

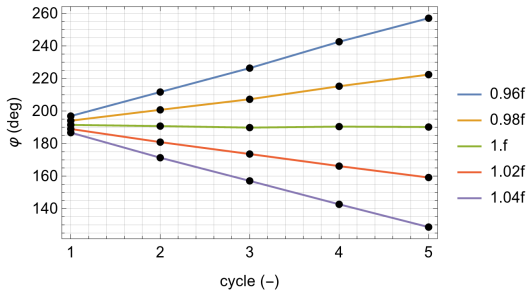


Figure 9: Phase angles of power drill's current without arcing

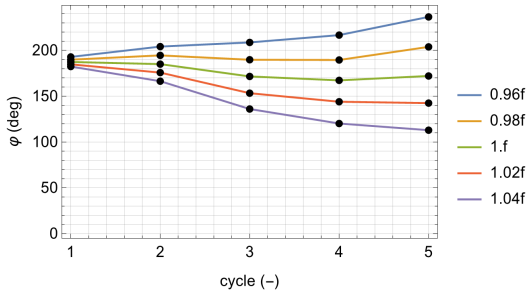
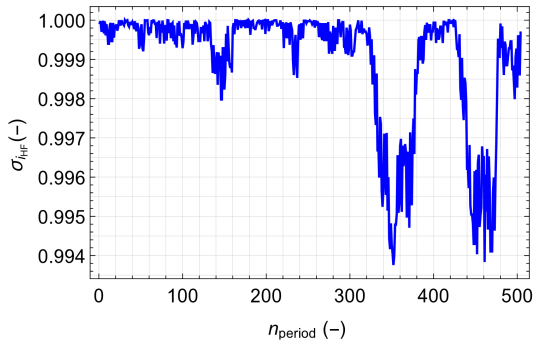


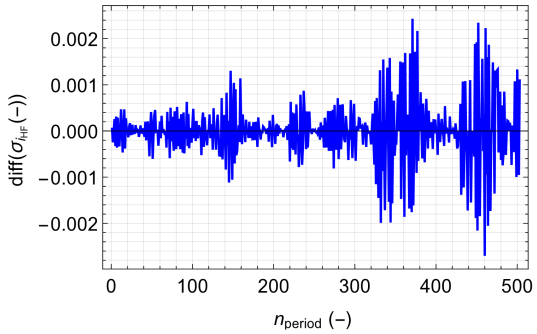
Figure 10: Phase angles of power drill's current with arcing

3.2. High frequency componetn of the current

The high frequency or hereafter i_{HF} components is a very known feature of a series arc. First of all, fundamental frequency of the analyzed signal should be taken into account. One of the significant characteristics of data is a standard deviation. It shows how far data lays from the mean, in other words, variability in the examined sample. Apparently, for a pure sinusoidal signal, it will be RMS value. For comparison of Standard deviation values of different signals is helpful to do normalization, so they can be compared in the same framework without losing information. The author considers that normalization by RMS value division is a viable approach. Moreover, in that case, normalized sigma or Standard deviation will be a dimensionless quantity.
For the signals analyzed hereafter, the author acquired high frequency components of the current signals using measurement setup. The fundamental frequencies were estimated, and standard deviations were calculated for the closest integer number of samples of the real value of the signals periods. Figure 11 shows characteristics of the signals without arcing. It is obvious that in those data is some variability and naturally differences of standard deviation show changes in that variability. Apparently, the main concern, in this case, is a source of the variability. Comparing this figure with figure of the signal with arcing 12, it becomes obvious that in case of the presence of a series arc the rate of change in variability is larger. Furthermore, it seems to be more random. The change shown in the figure 11b can be explained by the presence of noise. Consequently, from that observation, the hypothesis can be derived that infrequent changes in the variability of examined samples probably is an indicator of the presence of a series arc.

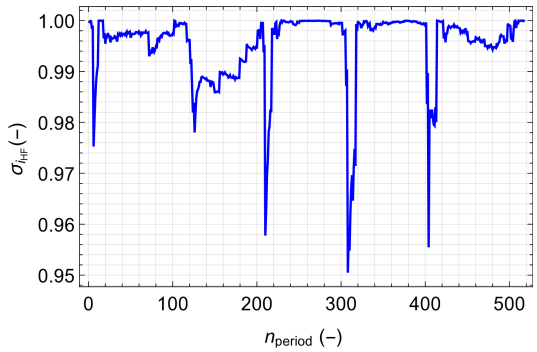


(a) Standard deviation

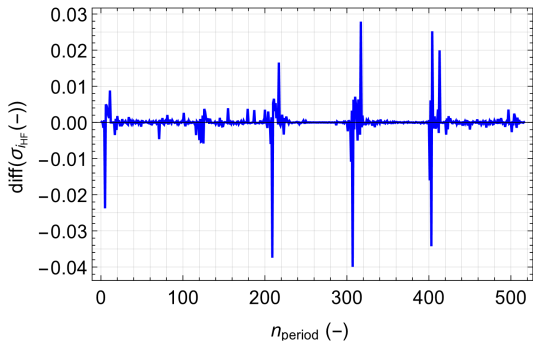


(b) Differences of standard deviation

Figure 11: i_{HF} of the power drill without arcing



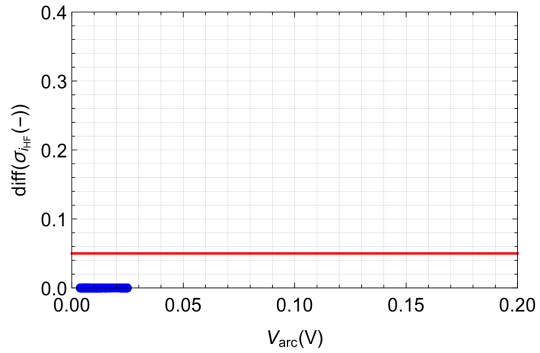
(a) Standard deviation



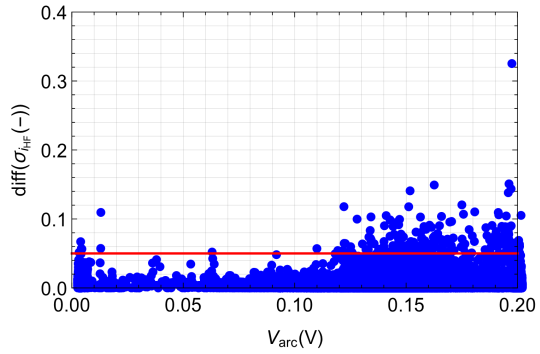
(b) Differences of standard deviation

Figure 12: i_{HF} of the power drill with arcing

The figure 13 shows the dependence between differences of SD and RMS of arc's voltage. Arc's voltage around zero volts corresponds to a short-circuited arc and around 0.2 volts to an open-circuit. That is clear that in case of arcing the values of $diff(\sigma_{i_{HF}})$ are higher. There is definitely a pattern, which, however, should be proven on a much bigger data set. Furthermore, there are exceptions that contradict author's hypothesis, some of high values of $diff(\sigma_{i_{HF}})$ could occur with a short-circuited arc 13b. These outliers are probably caused by switching. Therefore, in the future work more electric loads should be considered including their switching.



(a) Current without arcing

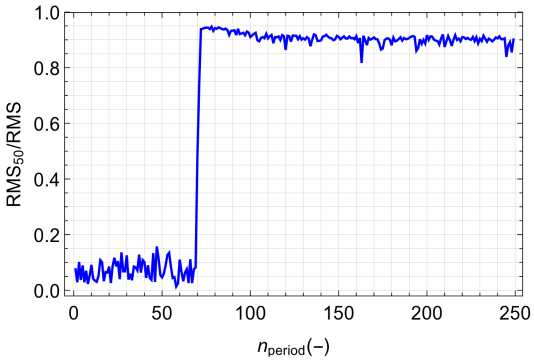


(b) Current with arcing

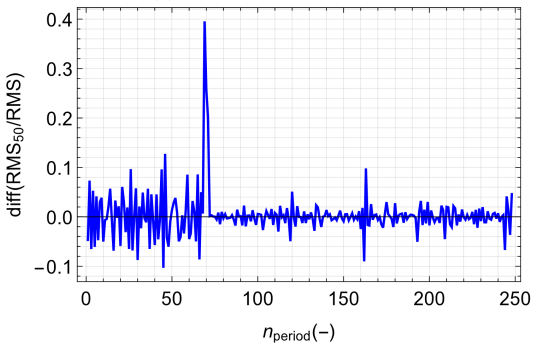
Figure 13: Differences of standard deviation of high frquency component of a drill's current

3.3. Fundamental harmonic and noise

The feature of a series arc to produce a broadband noise was investigated a lot. However, many approaches have time-consuming operations. On the contrary, faster algorithms, for instance, FFT, are not precise as was discussed in [8]. Hence, the author examined slightly different approach, which can give precision and does not require many calculations. The concept is similar to THD (Total Harmonic Distortion). Although, the author modified that to attempt to eliminate excessive operations without losing information about the signal. The steps are simple. Firstly, using method described in [9] estimate fundamental frequency. Secondly, calculate Fourier coefficients of the fundamental harmonic. Thirdly, compute RMS of the examined current signal. Finally, calculate the ratio of RMS of the fundamental harmonic to the RMS of the raw signal. Let us examine that with real data.



(a) RMS_{50} to RMS ratio



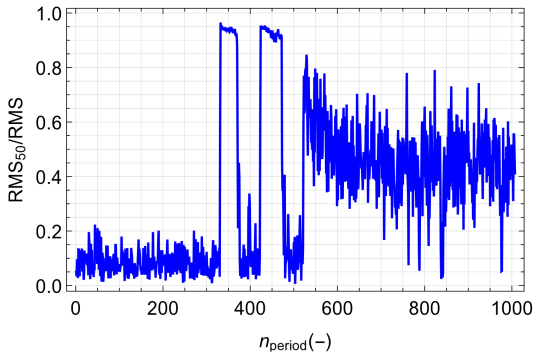
(b) Differences of RMS_{50} to RMS ratio

Figure 14: Drill's current without arcing

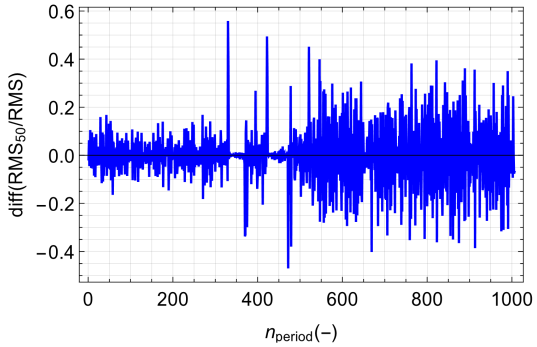
From figures 14 and 15 a conclusion can be derived that a change in RMS_{50} to RMS ratio could mean a change in load conditions or an arcing in a circuit. The small values of the ratio corresponding to the no-load operation. Furthermore, a change in the ratio higher than 0.4 seems to be an indicator of switching. However, these observations should be proven on a much high data set. The author observed on acquired data that RMS_{50} to RMS ratio higher than 0.2 and lower than 0.6 most likely correlate with an arcing condition. Figure 16 supports author's conclusions. The arc voltage lower than 0.02 is a short-circuited arc and arc voltage 0.2 is open circuit. There are certain outliers; however, the pattern is clear.

4. Consclusion

The paper presented another perspective on the series arc features. The investigation of phase angle during arc showed that the way the phase angle is computed is important as other factors. Therefore, any calculation involving Fourier Transform should consider grid frequency. Essentially, for the approaches outlines in the paper fundamental frequency is a very important input to the further analysis.

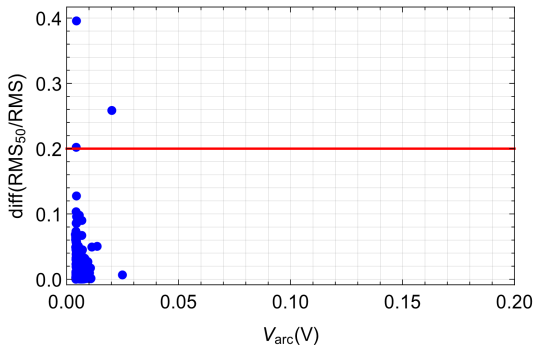


(a) RMS_{50} to RMS ratio

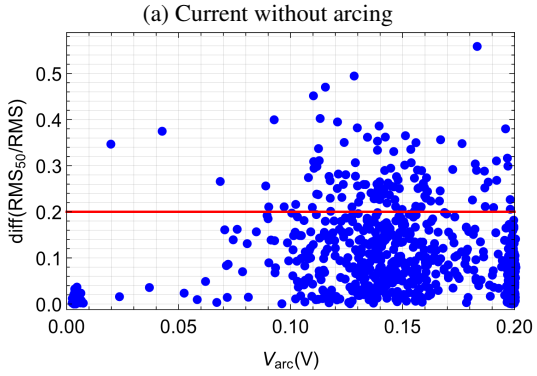


(b) Differences of RMS_{50} to RMS ratio

Figure 15: Drill's current with arcing



(a) Current without arcing



(b) Current with arcing

Figure 16: RMS_{50} to RMS ratio with corresponding RMS of arc voltage

Furthermore, the standard deviation analysis of the current signal is a promising indicator of the series arc fault. However, further research is definitely required to prove its for Arc Fault Circuit Interrupters. Moreover, the ration of the fundamental to noise can be used as well as the second indicator for AFCI. Hence, that will bring the redundancy to the detection algorithm. The future work will be focused on the acquisition of the larger data-set for analysis of the presented approaches.

Acknowledgment

Financial financial support of Czech Technical University in Prague through grant number by SGS17/181/OHK3/3T/13, is highly acknowledged.

References

[1] https://www.ishn.com/articles/102900-arcing-faults-cause-more-than-30000-home-fires-each-year. Accessed on: Apr. 27, 2018. [Online]

[2] Jones E, Oliphant E, Peterson P, et al. SciPy: Open Source Scientific Tools for Python, 2001-, `http://www.scipy.org/` Accessed on: Sep. 9, 2017. [Online].

[3] P. Muller, S. Tenbohlen, R. Maier and M. Anheuser, "Characteristics of Series and Parallel Low Current Arc Faults in the Time and Frequency Domain", 2010 Proceedings of the 56th IEEE Holm Conference on Electrical Contacts, Charleston, SC

[4] C. E. Restrepo, "Arc Fault Detection and Discrimination Methods", Electrical Contacts - 2007 Proceedings of the 53rd IEEE Holm Conference on Electrical Contacts, Pittsburgh, PA, 2007

[5] P. Muller, S. Tenbohlen, R. Maier and M. Anheuser, "Influence of Capacitive and Inductive Loads on the Detectability of Arc Faults", 2011 IEEE 57th Holm Conference on Electrical Contacts (Holm), Minneapolis, MN, 2011

[6] G. D. Gregory, Kon Wong and R. F. Dvorak, "More about arc-fault circuit interrupters", in IEEE Transactions on Industry Applications, vol. 40, no. 4, pp. 1006-1011, July-Aug. 2004.

[7] IEC62606, "General requirements for arc fault detection devices", 2013.

[8] L. Vetoshkin, "Methods of Detection of Serial Arc In the Presence of Household Electrical Loads", Praha, 2018. Diplomová práce (Ing.). České vysoké učení technické v Praze. Fakulta elektrotechnická, katedra elektroenergetiky. Vedoucí práce Jan Kyncl.

[9] J. Kyncl, A. Hariram and M. Novotny, "On measurement of synchronous phasors in electrical grids", 2013 IEEE International Symposium on Circuits and Systems (ISCAS2013), Beijing, 2013

[10] F. Frolov, "Určování frekvence a fázorů v elektroenergetice." Praha, 2016. Bakalářská práce (Bc.). České vysoké učení technické v Praze. Fakulta elektrotechnická, katedra elektroenergetiky. Vedoucí práce Jan Kyncl.

[11] A. F. Husain, "Series Arc Fault Detection in the Presence of Household Electrical Loads." Praha, 2017. Magisterská práce (Ing.). České vysoké učení technické v Praze. Fakulta elektrotechnická, katedra elektroenergetiky. Vedoucí práce Jan Kyncl.

[12] M. Vetterli, J. Kovačević, and V. K. Goyal, "Foundations of signal processing." Cambridge: Cambridge University Press, 2014.

[13] R. Shiavi, "Introduction to applied statistical signal analysis: guide to biomedical and electrical engineering applications", 3rd ed. Burlington: Academic Press, c2007.

[14] G. A. Hussain, "Methods for Arc-Flash Prediction in Medium Voltage and Low Voltage Switchgear.", Doctoral dissertation, Aalto University, 2015.

[15] Chapter 8. Spectrum Analysis. https://engineering.purdue.edu/ME365/Textbook/chapter8.png Accessed on Mar. 15, 2018. [Online]

About Author

Lavr Vetoshkin was born in 1995 in Tyumen, Russia. In 2013 he enrolled at the faculty of electrical engineering CTU. In 2016 he graduated with a bachelor degree in Applied Electrotechnics. He continued his studies at the master program and finally graduated in 2018 with a master degree in Power Systems. Currently, Lavr is a PhD student at the Department of Power Electrical Engineering at FEE CTU.

Effective calculation of heat flow trough walls

Jan Rimbala

[1]Department of Electrical Power Engineering, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

rimbajan@fel.cvut.cz

Abstract. *The paper deals with the possibility of replacing the partial differential equation solution of ordinary differential equations in determining the heat and electrical energy demand for heating buildings by heat pumps. After a theoretical description of the heat exchange model and the heat pump system, a case study is presented to analyze three typical building wall configurations: very well insulated, average insulated and non-insulated brick walls.*

Keywords

Heat transfer in solids, PDE, ODE, Heat Pump Heating of Buildings.

1. Functional description of heat pump

A heat pump is a set of specific components that together form a closed circuit. The individual parts could be divided into groups according to their basic function: a) The first group is heat exchangers. The heat exchanger serves to exchange energy between individual systems. Used heat exchangers are evaporator and condenser. b) The second group is the pressure part adjusting the pressure of the respective working medium. The pressure part includes a compressor and an expansion valve. The heat pump circuit is filled with the working substance. This is a special liquid with an extremely low boiling point. The evaporator is located outside a residential building in the place of a natural source called low-potential heat energy. The working medium enters the evaporator. The medium has low temperature and pressure. In the evaporator, the low-potential source transmits heat to the working medium according to the second law of thermodynamics. Since the working medium at a given pressure has a low boiling point, the supplied heat energy causes its evaporation. Another part of the circuit is the compressor. The input to the compressor is steam of the working medium and electric power for the drive. Steam is adiabatically compressed, the result will be the desired energy pumping from lower to higher. Another part of the circuit is a condenser that transmits energy to the heating system. During the heat transfer process, the working medium is cooled and then condensed. The liquid working medium comes out of the condenser and has a relatively high pressure. We lower the pressure of the medium isoenthalpy by means of an expansion valve. We lower the pressure of the medium during isenthalpic process by means of an expansion valve. The whole process is repeated.

2. Coefficient of performance

Coefficient of performance is the basic indicator of the heat pumps energy efficiency and ranks among the most important values. It is usually signed as COP or ϵ. The COP indicates the ratio of the heating power and the electrical input of the compressor. The COP value is a dimensionless quantity that is always greater than one.

$$COP = \frac{Q_{heat}}{P_{el}}[-] \tag{1}$$

The COP is related with temperatures of the source of low potential energy and water in heating systems. Thus we can estimate size of COP by:

$$COP_{carnot} = 0.5\frac{T_1 + 273.15}{T_1 - T_a} \tag{2}$$

where T_2 is temperature of water in interior and $T_¦$ is outside temeprature, both in Kelvin. We can see that this formula is basically reverse Carnot cycle. Constant 0,5 describes how are heat pumps efficient at this moment. Also the higher temperature of source and water temperature difference in the heating system, the smaller the heating factor.

3. Standard and proposed method of calculating heat flow through walls

Mathematical model of the system for the thermal field in the wall, the heat equation applies:

$$\rho c\frac{\partial T}{\partial t} = \nabla \cdot (\lambda \nabla T) + Q_v \tag{3}$$

$\rho\ (kg \cdot m^{-3})$ denotes the density, $c\ (J \cdot kg^{-1} \cdot K^{-1})$ the specific heat capacity, the $\lambda\ (W \cdot m^{-1} \cdot K^{-1})$ heat conductivity coefficient, and $Qv\ (W \cdot m^{-3})$ the bulk density of the internal heat sources. For a heat insulated wall we can

reasonably believe that the temperature in the space changes only in time and direction x and the internal heat sources are zero power. When considering these simplifications, the thermal equation has the form:

$$\rho c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}(\lambda \nabla T) + Q_v \tag{4}$$

The temperature field can be solved as a set of two partial differential equations supplemented by conditions of continuity of temperature and the heat flux vector on the contact of the thermal insulation and the original masonry. Another way of solving the temperature field is that we consider material properties as a function of time:

$$\rho = \rho(x);\ c = c(x);\ \lambda = \lambda(x) \tag{5}$$

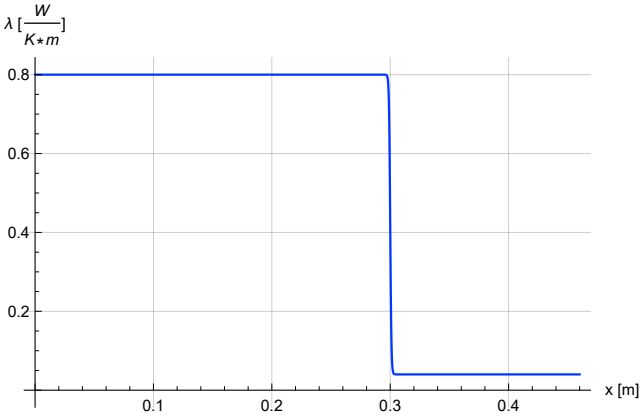


Fig. 1. Thermal conductivity in the wall

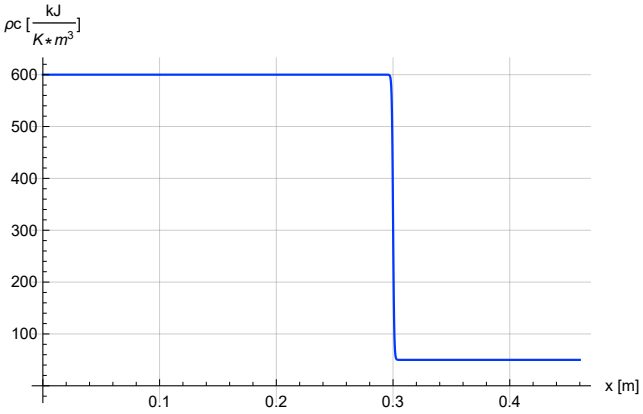


Fig. 2. Volume specefic heat capacity

The resulting shape of the heat transfer equation in the wall is:

$$\rho(x)c(x)\frac{\partial T(x,t)}{\partial t} = \frac{\partial T(x,t)}{\partial t}(\lambda(x)\frac{\partial T(x,t)}{\partial t}) \tag{6}$$

Let us consider the principle connection according to Fig. 4. T_A is the actual ambient air temperature, P_{HP} is the heat output of the heat pump, T_{in} temperature in the interior, T_1 the outlet water temperature, T_2 is the water temperature coming from the heating system. P_{BS} is the power of an additional heat source. In the case of monovalent operation, this power is zero. In case of bivalent operation, this power is turned on in situations where the heat pump output is not sufficient to cover the heat losses of the building. $P_{\sum}$ is the total power delivered to the interior. P_{el} is the electrical power of the heat pump. $\dot{m}$ is the mass flow rate of water in the heating system and c is the specific heat capacity of the water.

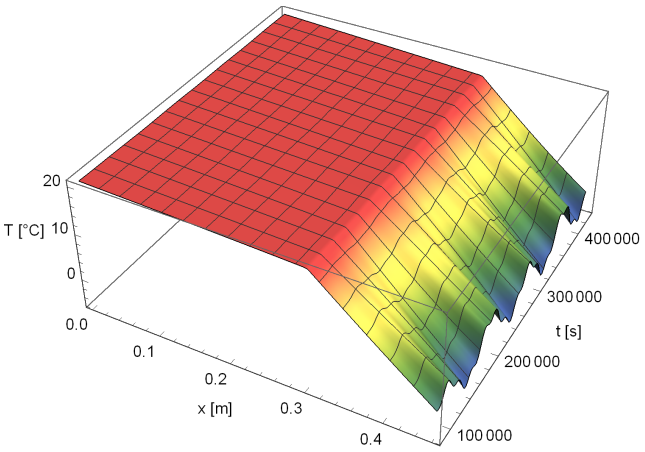


Fig. 3. Example of a solution of the temperature field in the wall

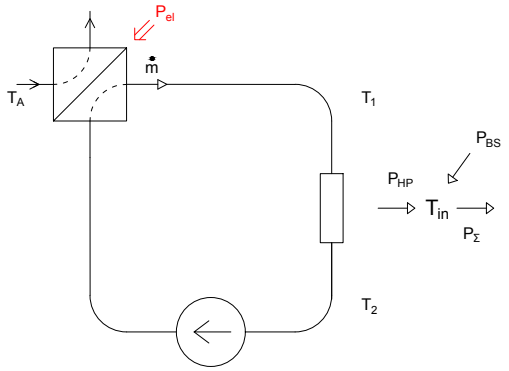


Fig. 4. Principal scheme

$$P_{HP} = \dot{m}c(T_1 - T_2) \tag{7a}$$

$$P_{HP} = COP(T_A, T_1) \cdot P_{el} \tag{7b}$$

$$P_{HP} = k \cdot (\frac{T_1 + T_2}{2} - T_{in}) \tag{7c}$$

$$P_{\sum} = P_{HP} + P_{BS} \tag{7d}$$

The calculation of the heat output passing through the walls by solving the partial differential equation is the most accurate and correct, but has several disadvantages.

One of them is the time-consuming calculation, when using common SW such as Wolfram Mathematica, or Agros2D, ANSYS and the like is quite considerable and the calculation over the whole heating season lasts even tens of minutes, while the calculation often occurs due to memory congestion computers.

The second disadvantage is that, when dealing with heat transfer in residential buildings, it is necessary to simultaneously address not only the passage of heat through the walls, which is described by the parabolic partial differential equation, but also the phenomena described by ordinary differential equations (eg. accumulation tank, air node) and nonlinear algebraic equations heat transfer through the windows, etc.).

E.g. SW Wolfram Mathematica does not allow simple solution of systems of simultaneous partial and ordinary differential equations, but systems of ordinary differential and algebraic equations are not a problem.

Logically, therefore, the idea is to try to replace the solution of partial differential equation by solving an ordinary differential equation, or a system of more ordinary differential equations. Solving parabolic differential equations using ODE is a common method, such as the popular Mthod of Lines, where equations are obtained by discretizing the PDE in a spatial coordinate leaving continuous time.

However, the Method of Lines requires fairly fine division and hence more ODE for sufficient accuracy.

The proposed method considers that a PDE solution is available for a certain time interval. This solution can be used to identify a substitute linear dynamic system.

To create a model, follow these steps (example 2nd order model):

- In a suitable program (for example, the author used SW Wolfram Mathematica) we solve the PDE describing the temperature field.
- We create tuples 8
- Using the least squares method we find the optimal linear fit in the form of 9

Initial conditions must be selected similarly to PDE solutions. We choose 10 Fortunately, heat transfer, as a dissipative system, ”forgets” the initial conditions. All we have to do is start the solution a little before the time interval being examined and use up part of the solution after the transient initial part.

$$\{q(t_i), \dot{q}(t_i), \ddot{q}(t_i), T_{in}(t_i) - T_{ext}(t_i)\} \quad i = 1, 2, \ldots, n \tag{8}$$

$$a_0 \cdot q(t) + a_1 \dot{q}(t) - a_2 \cdot \ddot{q}(t) = T_{in}(t) - T_{ext}(t) \tag{9}$$

$$q(0) = 0,\ \dot{q}(0) = 0 \tag{10}$$

4. Case study

Three cases were solved: brick wall 30 cm thick with thermal insulation made of expanded polystyrene 16 cm, brick wall 30 cm thick with thermal insulation made of polystyrene foam 8 cm and brick wall 45 cm thick without thermal insulation. Buildings with a maximum heat loss through walls of 6000W and 30 days of heating season in Brno 2012 was considered.

Results for brick wall 30 cm thick with thermal insulation made of expanded polystyrene 16 cm: The following graph shows the time dependence of specific heat flux obtained by PDE and ODE:

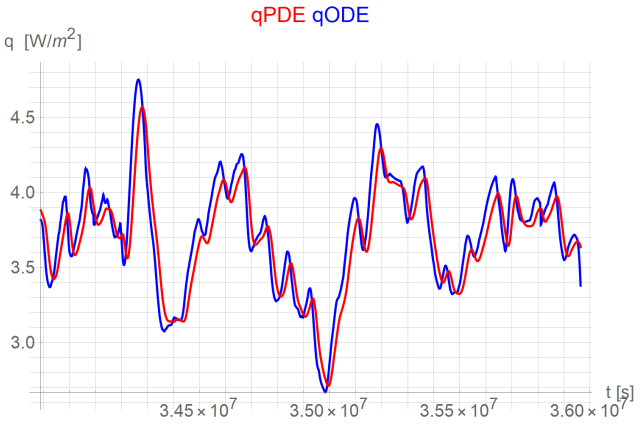


Fig. 5. Heat flow density bricks 30cm thermal isolation 16cm

The resulting linear fit of the internal and external temperature difference was: 4.59q[t]+18960q’[t]

The error of identifying the required electricity for the heat pump was 3.9 % and the total energy error needed to cover the heat loss of the wall was 2.8 %.

Results for brick wall 30 cm thick with thermal insulation made of expanded polystyrene 8 cm: The following graph shows the time dependence of specific heat flux obtained by PDE and ODE:

The resulting linear fit of the internal and external temperature difference was: 2.6q[t]+120993q’ [t]

The error of identifying the required electricity for the heat pump was -0.04% and the total energy error needed to cover the heat loss of the wall was 0.06%.

Results for brick wall 45 cm thick without thermal insulation: The following graph shows the time dependence of specific heat flux obtained by PDE and ODE:

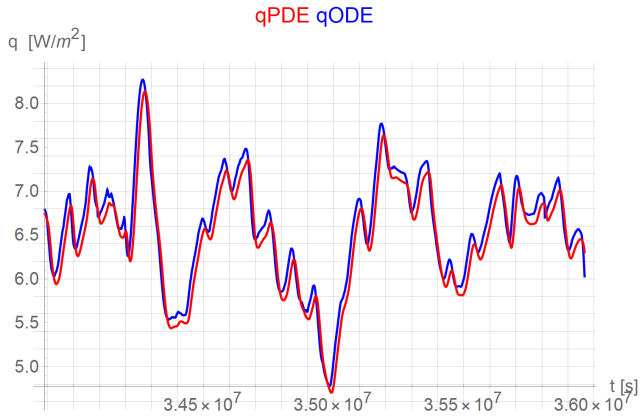


Fig. 6. Heat flow density bricks 30cm thermal isolation 8cm

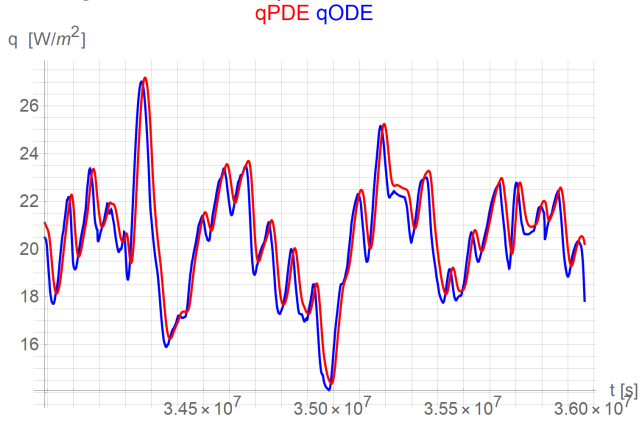


Fig. 7. Heat flow density bricks 45cm without thermal isolation

The resulting linear fit of the internal and external temperature difference was: 0.84q[t]+24410q' [t].

The error of identifying the required electricity for the heat pump was 0.93% and the total energy error needed to cover the heat loss of the wall was 0.82%.

Conclusion

Analyzes were carried out as in the given case study for models up to the fourth order. It has been shown that the use of higher order does not bring better accuracy in determining the required electrical energy or the energy needed to heat a building. The reason for this may be the high thermal inertia of the walls, which causes the higher derivatives to have very small amplitudes and thus the identification is inaccurate. This is evidenced by the large coefficients of higher order derivatives (up to fourteen orders of magnitude than the zero derivative). The computational speed using the proposed models is several orders of magnitude higher than when solving a partial differential equation and the model is easy to implement eg into MATLAB Similink. However, the first-order model gives a fairly sufficient accuracy of calculation when considering that standardized calculation procedures in the Czech Republic are even 100% different from the measurement of energy needs. Another direction of the research will be the use of similarity theory with the goal of finding correlation relations for the coefficients of the surrogate model based on the geometric and fabric properties of the walls.

Acknowledgements

Financial financial support of Czech Technical University in Prague through grant numberby SGS17/181/OHK3/3T/13, are highly acknowledged.

References

[1] H. Tsai, C. Hsu and C. Yang, "Design and performance evaluation of building integrated PVT and heat pump water heating (BIPVT/HPWH) system," 2013 IEEE 39th Photovoltaic Specialists Conference (PVSC), Tampa, FL, 2013, pp. 0821-0823.

[2] H. L. Tsai, C. Y. Hsu and B. Z. Lin, "Design and Evaluation of Photovoltaic/Thermal-Assisted Heat Pump Water Heating System," 2014 International Symposium on Computer, Consumer and Control, Taichung, 2014, pp. 59-62.

[3] R. Chang, Q. d. Yu and N. Zhu, "Part-Load Ratio Research on Energy Consumption of Heat Pump System," 2012 Asia-Pacific Power and Energy Engineering Conference, Shanghai, 2012, pp. 1-5.

[4] X. Chen, X. Zhang and H. Zhao, "Air source heat pump energy storage heating system for smart building," 2016 Chinese Control and Decision Conference (CCDC), Yinchuan, 2016, pp.

[5] J. H. Lienhard, J. H. Lienhard, "A heat transfer textbook," 4th ed. Mineola, N.Y.: Dover Publications, c2011. ISBN 978-0486479316.

[6] J. Ch. Hadorn, "Solar and heat pump systems for residential buildings, " Berlin, Germany: Ernst, 2015. ISBN 978-3-433-03040-0.

[7] K. Brodowicz, T. Dyakowski, "Heat pumps," Boston: Butterworth-Heinemann, 1993. ISBN 0-7506-0611-8.

About Author

Jan Rimbala was born in 1993 in Prague. He graduated his master degree in 2018 at the Czech Technical University in Prague in Department of Electrical Power Engineering. Since 2019 he is postgraduated student in Department of Electrical Power Engineering at Czech Technical University in Prague.

Model of operation modes of a stand-alone photovoltaic systems taking into account the actual operating conditions

Timofei ABRAMOV

Dept. of Economy and Management in Power Engineering, Czech Technical University, Technicka 2, 166 27 Praha, Czech Republic

Abramtim@fel.cvut.cz

Abstract. *Solar photovoltaic technology is one of the most rapidly growing renewable sources of electricity that has practical application in various fields of human activity due to its high availability, huge potential and environmental compatibility. The main technical problem of stand-alone photovoltaic systems is the need to harmonize the modes of production and consumption of electricity, which requires a reliable definition of the energy balance with a high time discretization. The nonlinearity of the characteristics of the main elements of photovoltaic plants, as well as their dependence on external factors, many of which are stochastic in nature, determines the high complexity and relevance of the problem of determining their energy characteristics in real operating conditions. The objective of this work is simulating of stand-alone photovoltaic system taking into account real operating conditions. Model will be created in MatLab Simulink.*

Keywords

Renewable, PV system, simulation, MatLab, energy balance.

1. Review of solar energy.

Since its first application in space missions in 1958, solar photovoltaics technology has come a long way. In Germany, a breakthrough in costs was observed over the last years, following a decade of massive investment in research and deployment. New solar photovoltaic power plants in Germany today cost almost 80 percent less than those built several years ago. Solar photovoltaics is already today a low-cost renewable energy technology.

2. Photovoltaic system.

The main components of PV system are: solar battery, with a series-parallel set of photovoltaic cells; DC-DC converter, which is operated by maximum power point tracking controller; storage system is based on storage battery; inverter. Frequently, solar tracker is applied for efficiency increasing.

The main technical problem of stand-alone photovoltaic systems is the need to harmonize the modes of production and consumption of electricity, which requires a reliable definition of the energy balance with a high time discretization. It leads to necessity to predict generated power by solar battery, SOC of storage battery, its' current and voltage (hold it in safety range). In order to analyze ability of renewable system meets the load requirements and provide uninterrupted power supply, with a specific maximum power, computer simulation tools is the most available, cheap and convenient, because allow to change a structure of PV system to fulfill the reliability requirements of power supply before installation.

3. Simulation of PV system.

Simulation of a stand-alone PV system includes the 7 main components: model of solar radiation, model of solar battery, daily temperature model, model of temperature of photovoltaic sell surface, storage battery, converters and load.

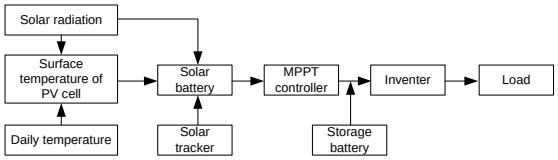


Fig. 1. Simulated components

In current article all components are simulated in MatLab/Simulink like subsystems, that allows to construct any configuration of PV station. Step of the simulation is 1 second and duration is 24 hours. The energy balance could be evaluated during one day of operating. Simulation of each component describes by following equations.

Value of primary available solar energy for conversion in PV system defines by intensity of solar radiation in the specific place and depends on geographical coordinates of installation. Total solar intensity on the arbitrary oriented surface is calculated by methodology offered by B.Y.H. Liu and R.C. Jordan [1].

$$I_{total}^{\beta\gamma} = I_{beam}^{\beta\gamma} + I_{dif}^{\beta\gamma} + I_{refl}^{\beta\gamma} =$$

$$= I_{beam}^{hor}\frac{\cos\theta}{\cos\theta_z} + I_{dif}^{hor}\cos^2\left(\frac{\beta}{2}\right) + \rho I_{total}^{hor}\sin^2\left(\frac{\beta}{2}\right)$$

θ – angle of drop, angle between the beam radiation on a surface and the normal to that surface; θ_z – zenith angle, the angle between the vertical and the line to the sun, that is, the angle of drop of beam radiation on a horizontal surface; β – slope, the angle between the plane of the surface and the horizontal; ρ – ground albedo. There are 8 inputs determine the value of the final solar intensity. The inputs are set via the parameter block Constant.

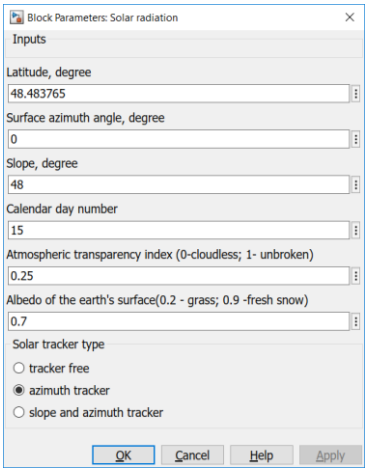


Fig. 2. Block parameters for solar radiation

In accordance with the process of changing the air temperature, it is obvious that the curve of the daily course of temperature has a form close to the sine wave.

$$T(t) = \overline{T} + 0{,}5\Delta T\cos\left(2\pi\frac{t_{cur} - t_{\max}}{t_{per}}\right)$$

$\overline{T}$ – daily average air temperature, ^oC;

ΔT – daily air temperature's amplitude, ^oC;

t_{per} –the period of air temperature changing, h;

$t_{\max}$ – local temperature maximum time, h;

t_{cur} – local sun time, h.

Solar battery effectiveness depends on surface temperature of PV cell significantly. As a result of operation, temperature of the solar panel can exceed the ambient temperature by 30ºC, so the surface temperature is quite important parameter. To calculate temperature, the empirical relationship is used:

$$\Delta T_{FM} = 0.0283G - 0.0058GV + 0.0005GV^2, ^oC$$

where G – level of solar radiation, W/m^2, V – wind speed, m/s.

Output power of solar battery is described by the equation:

$$P_{VV} = \frac{C_{FF}N_{FM}\eta_{con}G\ln\left(G\cdot 10^6\right)}{T_{FM}}$$

N_{FM} – number of photovoltaic modules; C_{FF} – solar battery ratio; η_{con} – effectiveness of converter with MPPT; $\frac{W}{m^2}$; T_{FM} -current surface temperature.

Output power depends of environment conditions. Physic stochastic values (wind speed, solar radiation, air temperature) are included in current surface temperature.

Solar battery ratio depends on photovoltaic module's parameters:

$$C_{FF} = \frac{FF\cdot T_{ref}\left(I_{sc} + k_I\left(T_{FM} - T_{ref}\right)\right)\left(V_{oc} + k_V\left(T_{FM} - T_{ref}\right)\right)}{G_{ref}\ln\left(G_{ref}\cdot 10^6\right)}$$

where, FF – filling factor of volt-ampere characteristic; T_{ref}, G_{ref} – temperature and illumination values of PV module under standard conditions; k_I, k_V – temperature's ratios of short-circuit current and open-circuit voltage. Filing factor describes nonlinearity of photovoltaic element.

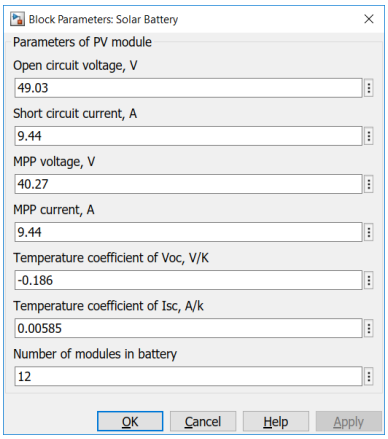


Fig. 3. Block parameters for PV module

Storage batterys' output voltage is described by equation offered by O.Treamblay and Louis-A. Dessaint [2]:

$$U_{batt} = E_0 - Ri - \frac{KQ}{Q - \int idt}\int idt + Ae^{-B\int idt}$$

E_0 – open-circuit voltage of cell, V; R – internal resistance, Ohm; i – current, A; K – polarization resistance, Ohm; Q – capacity, $A \cdot h$; $\int idt$ – current battery power, $A \cdot h$; A, B – coefficients characterizing the value of the voltage drop during the exponential discharge zone (V) and the inverse value of the cell capacitance at the end of the exponential discharge zone $\left(A \cdot h\right)^{-1}$.

Model of load is assigned by "Signal builder" that contains daily load plot of consumer.

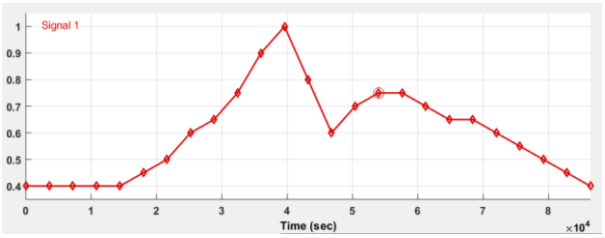


Fig. 4. Typical load plot for industry in settlement in the Signal Builder

For researching of energy characteristics were created "simplified" convertors, based on ideal sources of current and voltage.

All components of PV system have been simulated. Finally, its' view is following:

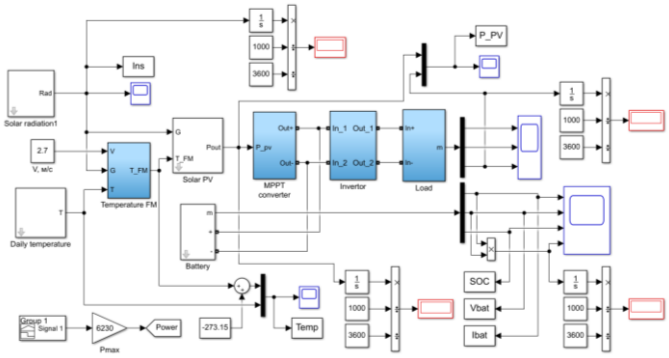


Fig. 5. PV system model

4. Results.

The simulation was provided for typical winter day (15th of January) for industry in settlement in the Far East in the Russia in Habarovskaya oblast' (Latitude 48.48^0 , Longitude $135{,}07^0$), maximum power of the load is 1500 W. Were chosen solar modules 72-cell mono PERC Double Glass Module JAM72D00-380/PR of the firm JA Solar. Was installed azimuth solar tracker, because it increases efficiency significantly, slope and azimuth tracker increase efficiency much lower in comparison with only azimuth tracker, but their price is higher.

The results of simulation are scopes of daily processes of solar radiation, daily temperature and surface temperature, output power and in storage battery (current, voltage, power and SOC).

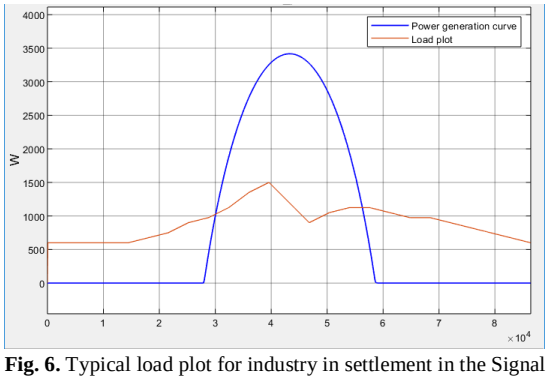


Fig. 6. Typical load plot for industry in settlement in the Signal Builder

Since sunrise until sunset the PV system generates power, and during exceeding the consumption value, the storage battery is charging and consumer is supplied by RES. At night only energy from storage battery is consumed. The rest consumption covers partially by PV system and partially by storage. The main issue is concerns SOC – is storage battery charging enough for night operating mode?

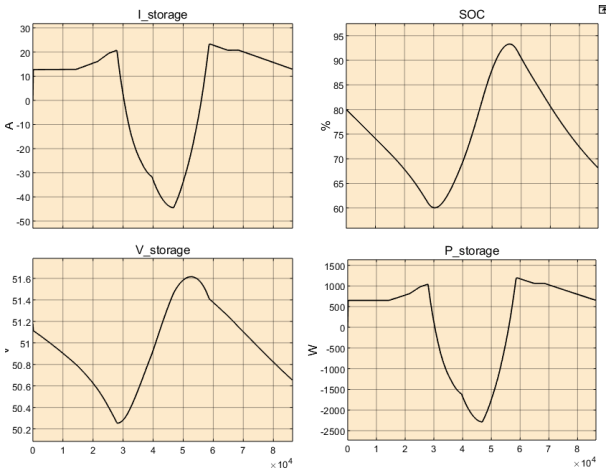


Fig. 7. Storage parameters

Four scopes describe storage parameters. During over generating the SOC is increasing, it leads to voltage increasing and negative power, that means income of power. Before sunrise initial value of SOC drops to 60%, but before sunset reaches 93% and then keep on falling to 68%. It means that consumed power hasn't restored until the initial value and that is the main motivation to change the parameters (change another solar modules or increase their number). Considering that this simulation has been provided for winter day, so the ambient conditions is the worst that could be.

And the scopes, that describes the quality of power energy for consumers.

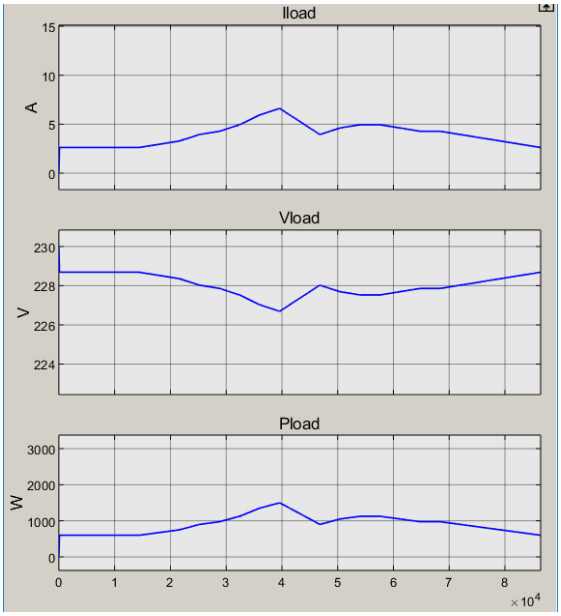


Fig. 8. Scopes for consumer

Voltage level doesn't drop under unacceptable value, so the quality requirements are fulfilled. And it is obviously, because exactly for this day capacity of the storage have been enough.

Conclusion

The submitted model of stand-alone photovoltaic system providing research of its work in real operating conditions. It takes into account the influence of external climatic factors on the energy characteristics of the solar battery, that increases the accuracy of the projected calculations in the generation of electricity for the specific region.

The results obtained in the simulation of power plant modes provide the possibility of a reasoned choice of the rational mode of operation of the power plant (year-round or seasonal), the required capacity of storage devices, the installed capacity and the nominal voltage of the solar battery, the type of solar tracker, the cross sections of conductors and the power of pre-converters, as well as the definition of effective algorithms for controlling the electric complex

Acknowledgments

Research described in the paper was supervised by full Prof. S. Obukhov, SEPE TPU in Tomsk and Prof. M.Benes FEL CTU.

References

[1] Liu B.Y.H., R.C.Jordan. Daily insolation on surfaces tilted towards the equator // ASHRAE Journal. – 1961. – V. 3. – P. 53–59.

[2] O. Tremblay. Experimental validation of a battery dynamic model for EV applications / O.Tremblay, L.-A. Dessaint // World electric vehicle journal. - 2009. - V.3. - P. 289 – 298.

About author

Timofei ABRAMOV was born in Yurga, Russian Federation in 1995. Graduated TPU, Institute of Power Engineering in 2017, bachelor of science, department of Power Supplying of Industrial Enterprises. Currently is a student of double degree master program between CTU and TPU.

Lightning Protection System for Power Supply Facility

Vladimir Konusarov[1]

[1] Dept. of Economics, Management and Humanities, Czech Technical University, Technická 1902/2, 166 27 Praha 6, Czech Republic

konusvla@fel.cvut.cz

Abstract. *The goal of this work is a development of optimal LPS for the indoor transformer substation 10/0,4 kV of city power distribution network. Lightning strikes and faults currents can pose a serious threat to living beings life and health. In addition to physical damage, these currents create risk of electrical systems failure and risk of fire. Designing technical reliable and cost effective LPS is highly important both for the protection of living beings and equipment. The solution of the problem is based on two steps implementation: design calculation and risk management end economical assessment. Risk management enables to assess protection measures and allows ascertaining if the protection system optimization is required. At the end of the paper, conclusions are made about the effectiveness of the developed LPS.*

Keywords

LPS, air-termination system, earth-termination system, down-conductor, earth rod, touch voltage, step voltage, damage, loss, risk, NPV.

1. Designing of LPS

LPS is one of the most important technological protection against lightning strikes and faults. Reliable protection of people and electrical equipment requires the creation of an effective LPS for the electrical installations. This work is devoted to the design and calculation of the LPS for the indoor transformer substation 10/0,4 kV of city power distribution network.

Protected substation is prefabricate type and consists of following main components:

- Two power transformers for 630 kVA;
- Medium and low voltage switchgears 10 and 0,4 kV respectively;
- Medium and low voltage interconnections;
- Auxiliary equipment (lightning, heating, signalization).

The LPS design process is determined by following steps based on information from [1], fig. 1. The process of designing can be iterative. It depends on location, architectural and aesthetic features of protected structure. An important part of designing is fulfilment of separation distance requirements that applies both internal and external LPS. Separation distances are determined by the protection level, the numbers of down conductors and position of the LPS. Sometimes, an alternative location or additional number of down conductors required to correspond the separation distance requirement [2].

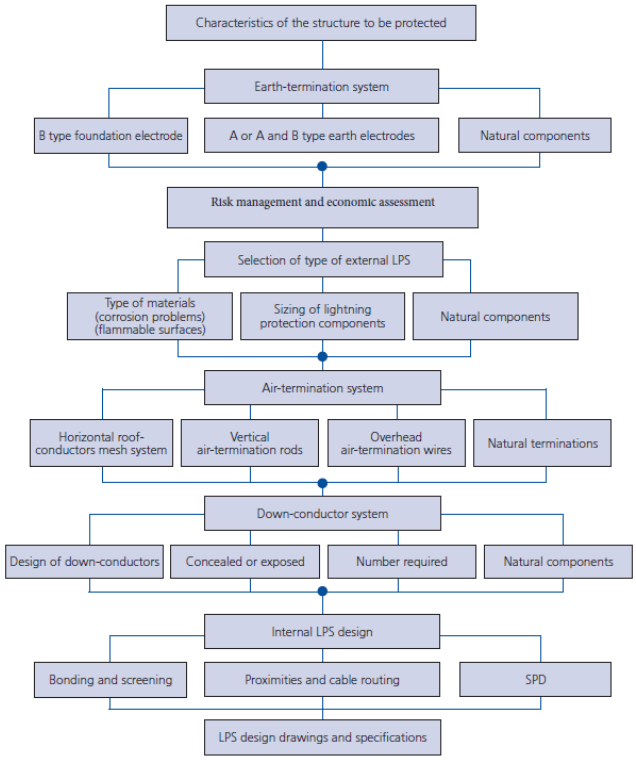


Fig. 1 LPS design process

As the protected structure has metal cover, air-termination system and down-conductors are not required here. A metal cover of the stepdown substation is also used as internal LPS for equipotential bounding. In this case only earth termination system should be designed for this type of substation [3].

1.1 Earth termination system calculation

In electrical installations with insulated neutral, the calculation of earth-termination system is performed by the following sequence:

The layout shown in fig. 2 will be taken as an earth-termination system layout.

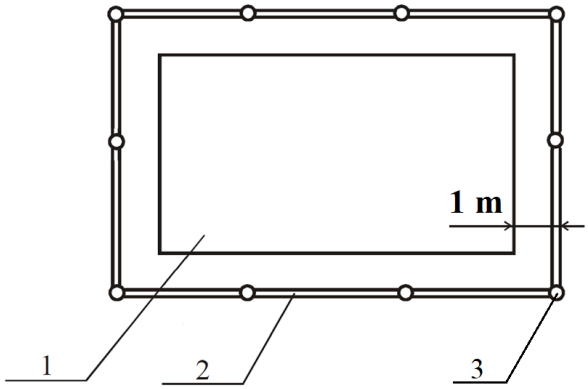


Fig. 2. Grounding device plan: 1 – the area occupied by installations (4,3×5,5) m ; 2 – horizontal conductor; 3 – vertical rod

It is taken to create the earth-termination system in the form of vertical rods with 5 m length and 12 mm cross section – l_{rod} = 5 m, d = 20 mm. The distance between the vertical rods is 2,3 m – a = 2,3 m, the primary number of rods is 12. Vertical rods are connected by horizontal steel tape with 40 mm width – b = 40 mm, the overall tape length can be found according to the grounding device plan:

$$l_{tape} = (4,3+2)\cdot 2 + (5,5+2)\cdot 2 \approx 27,6 \qquad (1)$$

The required resistance of grounding device is used both for electric installations with voltage up and higher of 1 kV should be not more than 4 Ohm [4]. In this case the grounding device resistance is taken equal to 4 Ohm – R_{GD} = 4 Ohm.

The soil resistivity is determined by the formula [3]:

$$\rho_{calc} = k_{seas}\rho \qquad (2)$$

where ρ is the soil resistivity measured for normal amount of moisture,
k_{seas} is the seasonality fluctuation coefficient that takes into cosideration the soil freezing and drying.

Soil at the substation for which is designed the grounding device is loamy. The ground resistivity is taken as ρ = 50 Ohm·m [5]. Seasonality fluctuation coefficient for the first climate zone, which is the zone with the most frigid climate, is taken as:
$K_{seasRod}$ = 1,9 is the seasonality fluctuation coefficient for the vertical ground conductor;
$K_{seasCond}$ = 5,8 is the seasonality fluctuation coefficient for the horizontal ground conductor;

The laying depth of the vertical ground conductor is f=0,7 m [5].

The vertical ground conductor resistance is determined by the formula [2]:

$$r_{rod} = \frac{\rho_{calc}}{2\pi l_{rod}}\left(\ln\frac{2l_{rod}}{d} + \frac{1}{2}\ln\frac{2t+\frac{l_{rod}}{2}}{2t-\frac{l_{rod}}{2}}\right) = 20,7\,Ohm \qquad (2)$$

where t is the depth from ground surface to the middle of rod,

$$t = f + \frac{l_e}{2} = 0,7 + \frac{3}{2} = 2,2\,m \qquad (3)$$

The ratio of the vertical ground conductor length to the distance between rods is equal to 0,5 – $a/l \approx 0,5$, utilization rate has the following values [3]:

η_{rod} = 0,52 – utilization rate for the vertical ground conductors;
η_{h} = 0,32 – utilization rate for the horizontal ground conductors;

The resistance of horizontal steel tape in grounding devise loop is determined by the formula [3]:

$$r_h = \frac{\rho_h}{\pi l_h}\ln\frac{1,5\cdot l_{tape}}{\sqrt{b\cdot t}} = 18,432\,Ohm \qquad (4)$$

The horizontal conductor resistance in the grounding device loop taking into account the utilization rate:

$$R_h = \frac{r_h}{\eta_h} = \frac{18,432}{0,32} = 57,6\,Ohm \qquad (5)$$

The required resistance of vertical rods is:

$$R'_{rod} = \frac{R_h\cdot R_{GS}}{R_h - R_{GS}} = \frac{57,6\cdot 4}{57,6-4} = 4,3\,Ohm \qquad (6)$$

The refined number of vertical conductors:

$$N'_{rod} = \frac{R_{rod}}{R'_{rod}} = \frac{41,4}{4,3} = 9,3 \to 10 \qquad (7)$$

Thus the clarified number of vertical ground conductors is 10. Two rods at the sides of the grounding device loop are not installed, fig. 2.

The vertical conductors resistance in this case:

$$R''_{rod} = \frac{r_{rod}}{N'_{rod}\cdot\eta'_{rod}} = \frac{21,5}{10\cdot 0,55} = 3,8\,Ohm \qquad (8)$$

The grounding device resistance will be determined as [3]:

$$R_{GS} = \frac{R_h\cdot R''_{rod}}{R_h + R''_{rod}} = \frac{57,6\cdot 3,8}{57,6+3,8} = 3,5\,Ohm \qquad (9)$$

The calculated grounding device resistance is less than standard value – 4 Ohm, it means that the earth-termination system is effective [6].

In this section the designing of the earth-termination system for the considered substation was performed. By choosing the preliminary size and number of electrodes the grounding device resistance was calculated, the result of calculation satisfies the requirements of the IEC standards. The given analytical calculations enable to make a risk management and economical assessment.

2. Risk management

For the beginning of risks evaluation, it is important to define the sources of damage, types of damage and types of losses for considered structure, tab. 1 [2].

Source of damage (Point of strike)	Type of damage	Type of loss
S1 Lightning flash to the structure	**D1** – Physical damage	$L1._A$ – Loss of human life $L3._A$ – Economic loss
	D2 – Damage due to fire and explosions	$L1._B$ – Loss of human life $L2._B$ – Loss of public service $L3._B$ – Economic loss
	D3 – Failure of systems	$L1._C$ – Loss of human life $L2._C$ – Loss of public service $L3._C$ – Economic loss
S2 Lightning strike near the structure	**D3** – Failure of systems	$L1._M$ – Loss of human life $L2._M$ – Loss of public service $L3._M$ – Economic loss
S3 Lightning strike to incoming line	**D1** – Physical damage	$L1._U$ – Loss of human life $L3._U$ – Economic loss
	D2 – Damage due to fire and explosions	$L1._V$ – Loss of human life $L2._V$ – Loss of public service $L3._V$ – Economic loss
	D3 – Failure of systems	$L1._W$ – Loss of human life $L2._W$ – Loss of public service $L3._W$ – Economic loss
S4 Lightning strike near incoming line	**D3** – Failure of systems	$L1._Z$ – Loss of human life $L2._Z$ – Loss of public service $L3._Z$ – Economic loss

Tab. 1. Parameters related to risk assessment

Risk assessment is conducted by steps shown in the fig. 3.

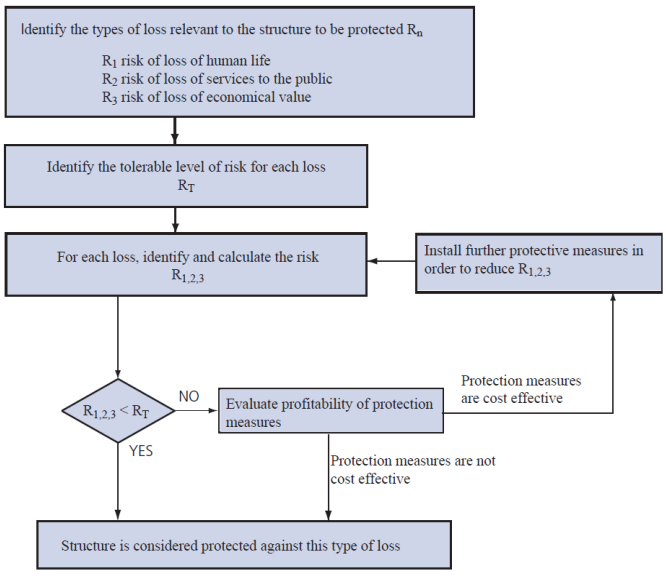


Fig. 3 Risk management steps

The risk $R_{1\text{-}3}$ that the lightning damage occurs is the sum of risks components. Risk components can be calculated by the following formula:

$$R_x = N_x \cdot P_x \cdot L_x \qquad (10)$$

where R_x –individual risk component,
N_x – number of dangerous events (the frequency of lightning strikes in a year),
P_x – probability of damage,
L_x – loss factor (quantity evaluation of damage).

The parameters of individual risk components are defined by protected structure features and shown in tab. 2.

Designation	Parameter description
Frequency of dangerous incidents caused by lightning strikes	
$N_D = 0{,}004$	to the structure
$N_M = 3$	near the structure
$N_L = 0{,}003$	to the incoming line
$N_I = 0{,}32$	near the incoming line
Probability of damage caused by lightning strike	
$P_A = 0{,}01$	Physical damage
$P_B = 1$	Accident damage (fire, explosions)
$P_C = 1$	Failure of internal systems
Probabilities of damage in case of nearby lightning strikes	
$P_M = 0$	
Probabilities of damage in case of a direct lightning strike to the incoming line	
$P_U = 0{,}1$	Physical damage
$P_V = 1$	Accident damage (fire, explosions)
$P_W = 1$	Failure of internal systems
Probabilities of damage in case of indirect lightning strikes to the incoming line	
$P_Z = 1$	Failure of internal systems
Losses results from	
$L_{1A}=L_{1U}=9{,}132\times10^{-8}$ $L_{1B}=L_{1V}=1{,}826\times10^{-8}$ $L_{1C}=L_{1M}=L_{1W}=L_{1Z}=1{,}826\times10^{-8}$	Loss of human life L_1
$L_{2B}=L_{2V}=6{,}667\times10^{-8}$ $L_{2C}=L_{2M}=L_{2W}=L_{2Z}=0{,}003$	Loss of public services L_2
$L_{3A}=L_{3U}=0$ $L_{3B}=L_{3V}=0$ $L_{3C}=L_{3M}=L_{3W}=L_{3Z}=0{,}015$	Economic loss L_3

Tab. 2. Parameters related to risk assessment

The risks $R_{1\text{-}3}$ are calculated from the sum of the appropriate components.

Risk of loss of human life:

$$\begin{aligned} R_1 &= R_{1A} + R_{1B} + R_{1C} + R_{1M} + R_{1U} + R_{1V} + R_{1W} + R_{1Z} = \\ &= N_D \cdot P_A \cdot L_{1A} + N_D \cdot P_B \cdot L_{1B} + N_D \cdot P_C \cdot L_{1C} + \\ &+ N_M \cdot P_M \cdot L_{1M} + N_L \cdot P_U \cdot L_{1U} + N_L \cdot P_V \cdot L_{1V} + \\ &+ N_L \cdot P_W \cdot L_{1W} + N_I \cdot P_Z \cdot L_{1Z} = 2{,}991 \cdot 10^{-6}\, 1/\,year \end{aligned} \qquad (11)$$

Risk of loss of services to the public:

$$\begin{aligned} R_2 &= R_{2B} + R_{2C} + R_{2M} + R_{2V} + R_{2W} + R_{2Z} = \\ &= N_D \cdot P_B \cdot L_{2B} + N_D \cdot P_C \cdot L_{2C} + N_M \cdot P_M \cdot L_{2M} + \\ &N_L \cdot P_V \cdot L_{2V} + N_L \cdot P_W \cdot L_{2W} + N_I \cdot P_Z \cdot L_{2Z} = \\ &= 1{,}092 \cdot 10^{-3}\, 1/\,year \end{aligned} \qquad (12)$$

Risk of loss of economic value:

$$R_3 = R_{3A} + R_{3B} + R_{3C} + R_{3M} + R_{3U} + R_{3V} + R_{3W} + R_{3Z} = \\ = N_D \cdot P_A \cdot L_{3A} + N_D \cdot P_B \cdot L_{3B} + N_D \cdot P_C \cdot L_{3C} + \\ + N_M \cdot P_M \cdot L_{3M} + N_L \cdot P_U \cdot L_{3U} + N_L \cdot P_V \cdot L_{3V} + \\ + N_L \cdot P_W \cdot L_{3W} + N_I \cdot P_Z \cdot L_{3Z} = 4{,}914 \cdot 10^{-3} 1/\ year \quad (13)$$

Now this calculated values of risk should be compared with tolerable values set by standard, tab. 3 [7].

Type of loss	R_T, 1/year	R, 1/year
L1	10^{-5}	$2{,}991 \times 10^{-5}$
L2	10^{-3}	$1{,}092 \times 10^{-3}$
L3	10^{-3}	$4{,}914 \times 10^{-3}$

Tab.3 Typical tolerable risk values

According to obtain results calculated risk values R_2 and R_3 exceed admissible limits, thus it is necessary to make evaluation of profitability of protection measures.

The equivalent annual costs of the total loss in the structure for the distribution company, which owns the substation – C_L are calculated by the formula:

$$C_L = \left(\sum_{t=0}^{T=20} \frac{R_3 \cdot c_s \cdot S \cdot (1-d)^t}{(1-\inf)^t} \right) \cdot a_t = 20\, EUR \quad (14)$$

$$a_t = \frac{(1+\inf)^T \cdot \inf}{(1+\inf)^T - 1} = \frac{(1+0{,}05)^{20} \cdot 0{,}05}{(1+0{,}05)^{20} - 1} = 0{,}08 \quad (15)$$

where c_s – the total value of the structure EUR/m^2 (since the real value of c_s is not available, it can be supposed equal to 200 EUR/m^2, based on Russian market value of the equipment);

S – area of the protected structure, m^2;

d = 3% – degradation factor (assumes loss of value of the installed equipment);

a_t – equivalent annual annuity factor;

inf – inflation level in Russia.

T – substation lifetime.

The equivalent annual costs C_{PM} of the protection measures for the company can be calculated by the following equation [8]:

$$C_{PM} = NPV \cdot a_t = 245 \cdot 0{,}10 = 25\, EUR \quad (16)$$

$$NPV = \frac{C_P}{r-g} - \frac{C_P}{r-g} \cdot \frac{(1+g)^T}{(1+r)^T} = \\ \frac{20}{0{,}08-0{,}03} - \frac{20}{0{,}08-0{,}03} \cdot \frac{(1+0{,}03)^{20}}{(1+0{,}08)^{20}} = 245\, EUR \quad (17)$$

$$a_t = \frac{(1+r)^T \cdot r}{(1+r)^T - 1} = \frac{(1+0{,}08)^{20} \cdot 0{,}08}{(1+0{,}08)^{20} - 1} = 0{,}10 \quad (18)$$

where C_P – stands for the annual costs of protection measures for the company (includes maintenance costs and repair),

NPV – net present value of protection measures implementation,

a_t – equivalent annual annuity factor,

r – discount rate (for financing the protection measures),

g – escalation rate (includes inflation and protection degradation)

T – protection lifetime in years.

Protection makes economic sense if the annual saving S_M is positive:

$$S_M = C_L - C_{PM} = 20 - 25 = -5\, EUR \quad (19)$$

Annual saving is negative value, thus the optimal solution is to not install the LPS protection, because relative annual costs for protection are higher than possible economical loss value in case of system failure.

3. Conclusion

In this paper the LPS for the indoor transformer substation 10/0,4 kV of city power distribution network was calculated. Despite the fact that results of risk management assessment indicated the need to implement further protective measures, economical evaluation showed that the optimal solution is to not install the LPS. It should be noted that risk of loss of human life was lower than tolerable value, so the additional costs connected with personal insurance were not included into the assessment. Depending on the size and complexity of the structure and the internal systems, different protection measures can be taken. Thus, there are several possibilities to protect the structure. The profitability of protection measures can therefore be further examined, even if an economically optimal solution has not been found straightforward, conducting the calculations consequence, an economically optimal solution can and should be achieved.

References

[1] ERICO: Lightning Protection Handbook Designing To The IEC 62305 Series Of Lightning Protection Standards, 2009

[2] DEHN + SÖHNE: Lightning Protection Guide 3rd updated edition, DEHN + SÖHNE GmbH, 2014

[3] Kabyshev A. V. "Molniyezashchita elektroustanovok sistem elektrosnabzheniya", TPU Tomsk, 2006.

[4] Hatim, G. "Calculation of Grounding Resistance and Earth Surface Potential for Two Layer Model Soil", International Journal of Electrical Engineering & Technology (IJEET), 2012, 158.

[5] Dashkovsky, A. "Calculation of the Grounding Device", TPU, Tomsk, 2008.

[6] Protection against lightning Part 1: General principles, stan. IEC 62305–1:2006

[7] Protection against lightning – Part 2: Risk management, stan. IEC 62305–2:2006

[8] Brealey R. A., Myers S. C., Allen F.:Principles of Corporate Finance, 10th edition, McGraw-Hill, 2011

Evaluation of substrate cleaning process based on contact angle measurement

Jakub JIRSA[1]

[1] Dept. of Microelectronics, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

jirsaja6@fel.cvut.cz

Abstract. *Wettability is one of the key factors of substrate characterization for inkjet printing method due to its influencing on drop formation. Wettability is often quantified by the surface contact angle between the droplet and the substrate. In this paper, we study the effect of wettability of glass substrates and their impact on printing nanofluid silver ink. Several methods of substrate treatment and their effect on wettability are compared and evaluated. For precision measurement of contact angle, the contact angle goniometer was build and tested. It enables easy and precise characterization of printable substrates such as glass, polyethylene terephthalate (PET) or polyethylene naphthalene (PEN).*

Keywords

Inkjet printing, surface treatment, surface wettability, surface characterization

1. Introduction

Inkjet printing is one of the key methods in low-cost device manufacture. It is a technology that allows printing various inks on many substrates such as glass, PET or kapton. This technology is commonly used for electronic circuit and device creation. However, inkjet printing becomes challenging due to the high demand for printing resolution and miniaturization of devices.

One of the crucial features with high impact on printing resolution is surface wettability [1]. Surface wetting is associated with the behavior of printed droplets on substrates. The contact angle of the droplet can describe the wettability. Substrates with a contact angle lower than 90° are called hydrophilic, when contact angle is greater than 90° and less than 150° substrates are called hydrophobic and become superhydrophobic once the contact angle is greater than 150° [2]. The contact angle of the droplet can be controlled by the substrate cleaning process, modification of ink or printing process. The most common methods for substrate cleaning are: wet methods (alcohol based), ozone and plasma treatment [3-4].

In this work, several cleaning processes of glass substrates are compared and evaluated based on the contact angle with a substrate. Printing of nanoparticle silver ink is then done and evaluated.

2. Experiment details

2.1 Cleaning process

Cleaning substrate is essential for deposition methods such as spin coating, atomic layer deposition or inkjet printing. Here three-step method is used and contact angle is evaluated after each step.

Firstly glass substrate is ultrasonicated in deionized water at 50°C with Hellmanex III (100:1) for 10 minutes. The sample is then ultrasonicated at 50°C in isopropyl alcohol for 10 minutes. Finally, to finish the cleaning process and ensure the hydrophilic substrate surface, the substrate is put in a UV ozone cleaner for 15 minutes.

After each step one substrate is put out of the process and 10 pl droplets of water and glycol are put on the substrate. A contact angle of each droplet is measured with contact angle goniometer.

2.2 Contact angle measurement

The contact angle goniometer consists of three parts: high-resolution camera, table for substrate and backlight. The substrate is laid on the measurement table, recording of the videosequence is started and a droplet is put on the substrate. The recorded sequence is then analyzed and contact angle is calculated from its frames.

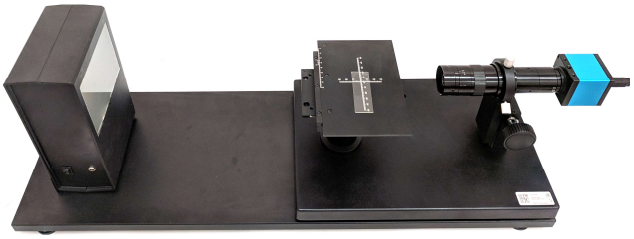


Fig. 1. The contact angle goniometer

3. Results and discussion

The cleaning process described in section 2.1 was used to clean four glass substrates. After each step one of the substrates was removed of the cleaning process, water and glycol droplets were put on the substrate and contact angles were measured. The impact of each procedure on wetting is shown in Fig. 2.

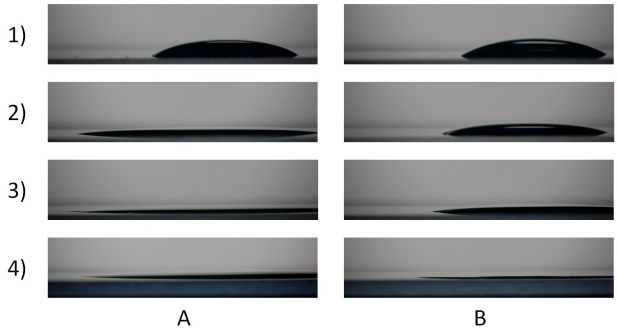


Fig. 2. Impact of the glass substrate cleaning process on the shape of water (A) and glycol (B) droplets. Numbers mark the phase of substrate cleaning: 1 – no treatment, 2 – hellmanex III, 3 – isopropanol, 4 – UV ozone cleaner.

Contact angles measured after each cleaning step are shown in Tab. 1

	Deionized water	Glycol
No treatment	30.6°	33.3°
Hellmanex III	7.6°	25.7°
isopropyl alcohol	< 5°	7.9°
UV Ozone	< 5°	< 5°

Tab. 1. Contact angles before, during and after cleaning

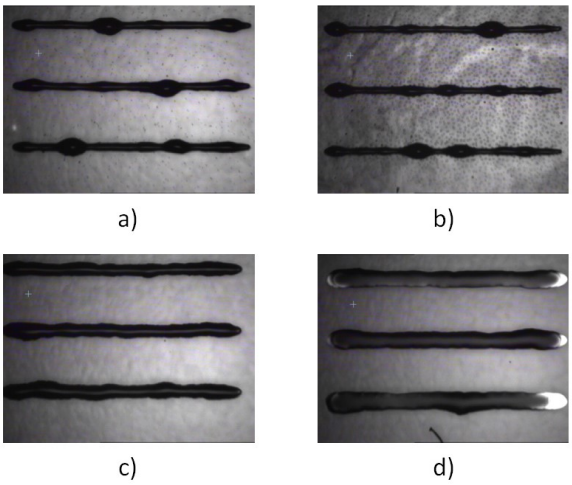


Fig. 3. Silver ink lines on a glass substrate printed with Dimatix DMP-2831 a) No treatment b) Hellmanex III c) Isopropyl alcohol d) UV ozone cleaner

Cleaning process has high impact on precision of printing as can be seen in Fig. 3. Printing on dirty or partially cleaned substrates leads to printing defects.

The study proves that (i) cleanness of the substrate can be evaluated based on contact angle measurement and (ii) substrate cleanness is the prerequisite for high resolution printing on glass.

Acknowledgements

This work was supported by the Czech Grant Agency of the Czech technical University in Prague, grant No. SGS17/188/OHK3/3T/13.

The author would like to thank doc. Ing. Václav Prajzler, Ph.D. and Ing. Alexand Laposa from Department of Microelectronics.

References

[1] SAEID, V., CHRISTOPHER, T., ASHCROFT, I., WILDMAN, R. Surface microstructuring to modify wettability for 3D printing of nano-filled inks. *Chemical Engineering Research and Design,* 2016, vol. 109, p. 414 – 420.

[2] PIETRIKOVA, A., LUKACS, P., JAKUBECZYOVA, D., BALLOKOVA, B., POTENCKI, J., TOMASZEWSKI, G., PEKAREK, J., PRIKRYLOVA, K., FIDES, M. Surface analysis of polymeric substrates used for inkjet printing technology., *Circuit World.,* 2016, vol. 42, no. 1., p. 9 – 16.

[3] LIN, J., DAHLSTEN, P., PEKKANEN, J., LINDÉN, M., MANTYSALO, M., OSTERBACKA, R. Surface energy patterning for inkjet printing in device fabrication., *Proceedings of SPIE.,* 2009.

[4] YEOL PARK, H., KANG, B., J., LEE, D., HOON, O. Control of surface wettability for inkjet printing by combining hydrophobic coating and plasma treatment., *Thin Solid Films.,* 2013, vol. 546, p. 162 – 166.

About Author

Jakub Jirsa received a Msc. in sensors and measurement from the Czech Technical University in Prague in 2018. He is currently working as a Ph.D. student at the Department of Microelectronics. His current work is focused on digital design of pixel detectors.

Implementation of Small Wind Turbine in Matlab Simulink

Jakub URBANSKÝ[1], Michal ŠPES[2], Daniel PÁL[3], Michal MÁRTON[4]

[1,2,3] Department of Electrical Power Engineering, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Mäsiarska 74, 041 20 Košice, Slovak Republic
[4] Department of Electronics and Multimedia Telecommunications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Vysokoškolská 4, 040 01 Košice, Slovak Republic

Jakub.Urbansky@tuke.sk, Michal.Spes@tuke.sk, Daniel.Pal@tuke.sk, Michal.Marton@tuke.sk

Abstract. *Nowadays we can notice the increasing interest in practical application of renewable energy sources around the world. The most common types are photovoltaics panels and wind turbines. They can be utilized in large farms but also in households in form of off-grid systems or as supplementary energy sources. Therefore, estimating their electricity production is important task. The main goal of this article is investigate theoretical background of small household wind turbines and create a computer simulation of one turbine energy production. In order to carry out needed simulations, the computer model of wind turbine was implemented in Matlab Simulink environment. Based on carried out simulations with usage of real wind speed measurement realized at department of electrical power engineering we can state potential of small wind turbine in household application.*

Keywords

Wind turbine, Implementation, Renewable energy sources, Computer simulation, Matlab Simulink.

1. Introduction

With the ever-increasing number of the world's population, the demand for energy is increasing, especially for electricity. It is widely known fact that the energy stored in fossil fuels will be depleted over the next few years. It is therefore important to develop new sources of energy and to deploy already existing alternative and renewable resources to meet the global demand for energy with the least possible burden on the environment.

At present, the production of electricity from renewable sources, both in large photovoltaic and wind power plants, as well as in local conditions such as production halls, large settlements, family houses, etc., is gaining popularity at both domestic and international levels.

In year 2018 gross annual wind installations in Europe decreased to 11.7 GW, with 0.4 GW being decommissioned (see Fig. 1.). It was a significant decrease compared to 17.1 GW in 2017 which was record year for Europe. [1]

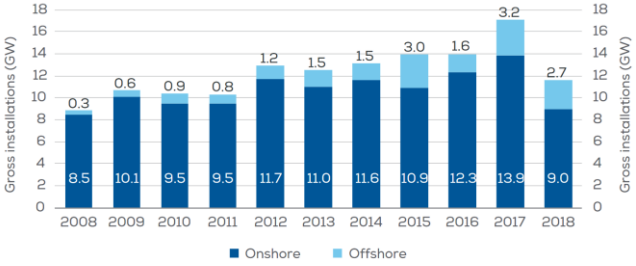


Fig. 1. Gross annual onshore and offshore wind installations in Europe [1]

The main aim of this article will be investigate possibilities of electricity generation from small wind turbine in locality of Košice.

The primary goal is to create model of the particular wind turbine in simulation software Matlab Simulink. The secondary goal will be carry out simulation based on real wind speed measurements realized on Department of Electrical Power Engineering.

2. Wind energy

Solar radiation heats both the Earth's surface and the atmosphere. The temperature difference between the atmosphere and the earth's surface causes air to flow that is the source of wind energy. [2]

The energy of moving air mass can be expressed by the kinetic energy equation (1) [3]:

$$E_k = \frac{1}{2}.m.v^2\,(J) \tag{1}$$

Where:

m - mass (kg)

v - moving mass velocity (m.s^{-1})

The following equation (2) applies to mass [3]:

$$m = \rho.V = \rho.A.s\,(kg) \tag{2}$$

Where:

A - area by which the given air volume flows (m^2)

s - distance which the moving air passes (m)

From (1) and (2) for wind power flowing through the unit area is possible to establish equation (3) [3]:

$$P_v = \frac{E}{A.t} = \frac{1}{2}.\rho.\frac{A.s}{A.t}.v^2 = \frac{1}{2}.\rho.v^3\,(W.m^{-2}) \tag{3}$$

2.1 Wind Turbine

A wind turbine is a device that converts the wind's kinetic energy into mechanical energy. In wind power plants, kinetic energy is converted into mechanical energy, which is then transformed into electrical energy. [3]

By axis rotation is possible to divide wind turbines as (see Fig. 2.) [3]:

- vertical,
- horizontal.

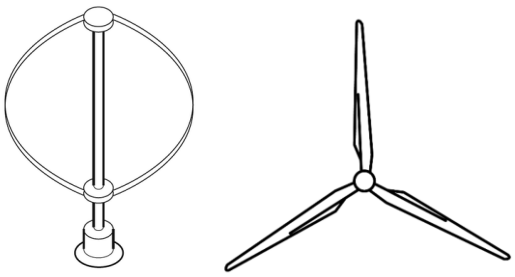


Fig. 2. Vertical (left), horizontal (right) turbine

3. Model description

Due to low average wind speeds in location of Košice, turbine, Airdolphin PRO Z-1000-48 (Fig. 3.) with low cut-in wind speed (see Tab. 1.) was chosen. [4]

Fig. 3. Chosen wind turbine [4]

Wind Turbine Type	Horizontal axis, up-wind
Rotor Diameter	1800 mm
Mass	17,5kg
Number of Blades	3
Cut-in Wind Speed	2,5 m.s^{-1}
Cut-out Wind Speed	-
Survival Wind Speed	65 m.s^{-1}
Rated Power	1kW at 12,5 m.s^{-1}
Rated Rotor Speed	1000 rpm
Maximum Power	2,3kW at 20 m.s^{-1}

Tab. 1. Technical specification of wind turbine [4]

Technical specifications of chosen wind turbine are listed in Tab. 1. Nominal power of wind turbine is 1kW at wind speed of 12,5 m.s^{-1}.

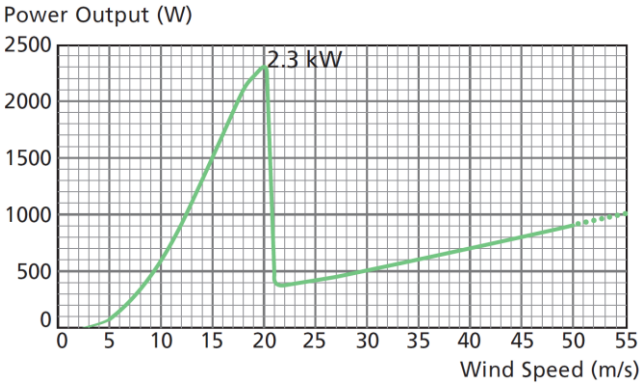


Fig. 4. Power output characteristic of wind turbine [5]

Power output characteristic of Z-1000-48 is shown at Fig. 4. with peak power of 2,3 kW at wind speed of 20 m.s^{-1}.

4. Implementation in Matlab Simulink

Wind turbine is implemented in Matlab Simulink based on power output characteristic shown on Fig. 4. Cut-in wind speed of turbine is 2,5 m.s^{-1} and survival wind speed is 65 m.s^{-1}, such high wind speed isn't common for the local climatic conditions. Thus, 55 m.s^{-1} is enough, and was set as cut-of wind speed in simulation.

$$P(v) = \begin{cases} P = P_{0\text{-}2,5}; & v < 2,5\ m.s^{-1} \quad (4) \\ P = P_{2,5\text{-}20}; & 2,5\ m.s^{-1} \le v < 21\ m.s^{-1} \quad (5) \\ P = P_{21\text{-}55}; & 21\ m.s^{-1} \le v < 55\ m.s^{-1} \quad (6) \end{cases}$$

For wind speeds 0 – 2,5 m.s^{-1} is power output based on Fig. 4. defined as 0W.

Therefore, for (4) can be mathematically stated as:

$$P_{0-2,5} = 0\,W \tag{7}$$

For wind speeds 2,5 – 20 m.s^{-1} is power output based on Fig. 4. defined as approximated curve with polynomial equation of fourth order.

Therefore, for (5) can be mathematically stated as:

$$P_{2,5-20} = -0,04v^4 + 1,33v^3 - 5,69v^2 + 30,92v - 65,88\,(W)\,(8)$$

With accuracy of $R^2 = 0,99$

For wind speeds 21 – 55 m.s^{-1} is power output based on Fig. 4. defined as approximated curve with polynomial equation of fourth order.

Therefore, for (6) can be mathematically stated as:

$$P_{21-55} = 0.0008v^4 - 0.1332v^3 + 7.9905v^2 - 187.79v + 1872.2\,(W)\,(9)$$

With accuracy of $R^2 = 0,96$

For wind speeds 55 – ∞ m.s^{-1} is power output set to 0 W.

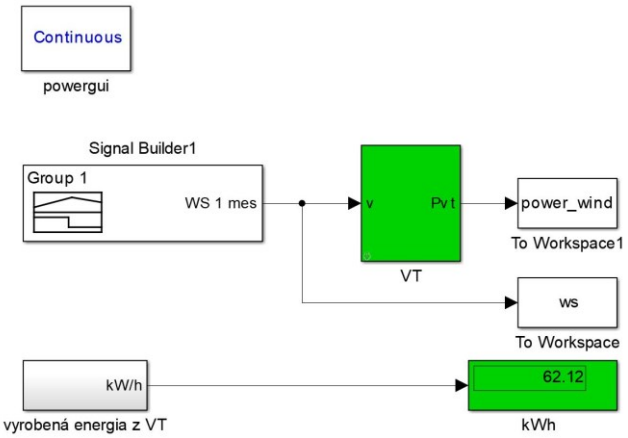


Fig. 5. Implementation of model in Simulink

At Fig. 5. is shown final model implemented in Matlab Simulink. Input parameter is wind speed in m.s^{-1} and output is in watts. Model is also calculating produced electric energy in kWh.

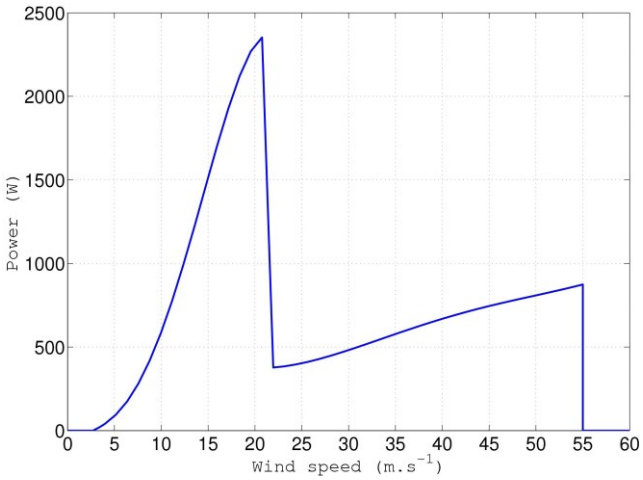


Fig. 6. Implementation of power output characteristic in Simulink model

Final implementation of power output characteristic based on (7), (8) and (9) is shown at Fig. 6.

5. Simulations

Simulation was running in continuous mode with ode45 solver which is capable of solving nonstiff differential equations — medium order method, with simulation step of 1s.

Input parameter was values of wind speed measured at Department of Power Engineering for period of one month. Specifically, month January of 2015 was selected, with average wind speed of 3,176 m.s^{-1}. Time course of measured wind speed is shown at Fig. 7. Based on model calculation wind turbine produced 62,12 kWh of electricity (Fig. 5.).

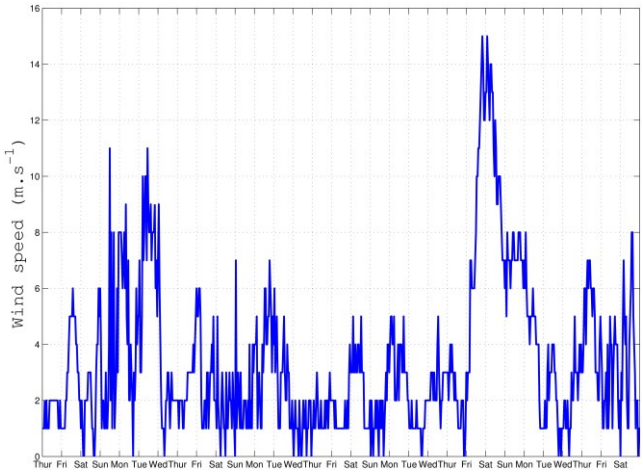


Fig. 7. Time course of measured wind speed for month January of 2015

Measuring sample time was 3600s (1h), thus it has to be corrected before inserting in Matlab. As simulation time in Matlab was 1s, more suitable sample time of wind speed measurements would be 1s as well to get as accurate results as possible.

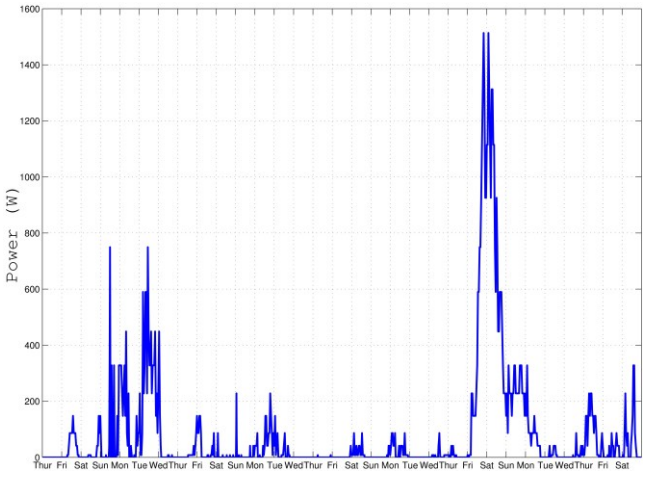


Fig. 8. Time course of calculated wind turbine power output for month January of 2015

At Fig. 8. time course of calculated wind turbine power output for month January of 2015 is shown. It is possible to conclude based on peak wind speeds for 22th and 23th from Fig. 6., that peak in wind electricity

production also match 22th and 23th day of January of year 2015.

Conclusion

At present time it is possible to notice the increasing interest in practical application of renewable energy sources around the world. Since the most common types of these energy sources are photovoltaics panels and wind turbines. We've decided to investigate possibility of electricity production from wind in local conditions.

The main goal of this article was to investigate theoretical background of small household wind turbines and create a computer simulation of one specific wind turbine energy production. Since most of the wind turbines have high cut-in wind speeds, turbine with one of lowest possible cut-in speed was chosen.

This turbine power out characteristic provided by manufactures was successfully implemented in to Matlab Simulink environment via polynomial approximation where we specified 2 stages with mathematical equations and 2 via constant 0 W value. Established model is dynamically calculating power output of wind turbine based of wind speed on input. It would be interesting to verify this modeling approach with real wind turbine and wind speed measuring sample time of 1s.

It is possible to conclude that such wind turbine in our local conditions wouldn't make sense. Produced electricity wasn't sufficient enough to justify purchase costs. Price of such turbine is circa 3600€, it is possible to buy similar wind turbines from China for fraction of cost, however with questionable durability and reliability. For that price would be in our local conditions much more reasonable to invest financial resources in to photovoltaic panels, circa 18 panels or more cloud be purchased for similar prica, and produce much more carbon free electricity this way.

Acknowledgements

This work was supported by the Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences by the projects VEGA No. 1/0372/18.

References

[1] Wind EUROPE: Wind energy in Europe in 2018, p. 11 [Online]. Available at: <https://windeurope.org/wp-content/uploads/files/about-wind/statistics/WindEurope-Annual-Statistics-2018.pdf> Accessed on March 10, 2019

[2] QUASCHING, V. Understanding Renewable Energy Systems in Earthscan. Earthscan, 2005, ISBN 1-84407-128-6, p. 19.

[3] MASTNÝ, P. Obnovitelné zdroje elektrické energie in Praha. České vysoké učení technické v Praze, 2011. ISBN 978-80-01-04937-2, p. 37.

[4] Manualslib: Zephyr Airdolphin Pro Z-1000 - 48 Instruction Manual [Online]. Available at: < https://www.manualslib.com/manual/846818/Zephyr-Airdolphin-Pro-Z-1000-48.html> Accessed on March 10, 2019

[5] e-Marine-Systems: Airdolphin 48V Pro Brochure [Online]. Available at: < http://www.emarineinc.com/pdf/WGZ30048.pdf> Accessed on March 10, 2019

About Authors...

Jakub URBANSKÝ was born in 1991. In 2017 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. At present is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in electric power engineering on subject of the integration of electric vehicle into the electrical power distribution, using the vehicles-to-grid (V2G) technology. His scientific research is mainly focused on research of renewables.

Michal ŠPES was born in 1991. In 2015 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. At present is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in electric power engineering on subject evaluation of generator exciting outage. His scientific research is mainly focused on research of powerline ampacity system.

Daniel PÁL was born in 1994. In 2018 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. At present is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in electric power engineering on subject of the design of the reconstruction of the school lightning. His scientific research is mainly focused on research of renewables.

Michal MÁRTON was born in 1992. In 2016 graduated (MSc) at the Department of Electronics and Multimedia telecommunications. At present is a Ph.D. student in the Department of Electronics and multimedia telecommunications on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in multimedia telecommunications on measurement with the fibre optic gyroscope system. His scientific research is mainly focused on research of optical communications.

Analysis of integration of accumulation systems in Smart Grid

Maksym Oliinyk[1], *Daniel Pál* [2]

[1] *Department of Electric Power Engineering; Faculty of Electrical Engineering and Informatics*
Technical University of Košice
Košice, Slovak Republic

[2] *Department of Electric Power Engineering; Faculty of Electrical Engineering and Informatics*
Technical University of Košice
Košice, Slovak Republic

Maksym.oliinyk@tuke.sk, daniel.pal@tuke.sk

Abstract: *This article focuses on modern energy storage systems. The importance of the further development of modern energy storage systems and above all large quantities of electricity is associated with the widespread introduction of renewable energy sources (RES) in the world in recent decades and exist serious problems in Western countries with providing high-quality continuous and break-even generation of electricity from them. The analysis of the benefits of introducing energy storage systems in a small electrical network. In addition, an analysis is conducted of the possibility of the so-called temporary arbitration of electricity prices, and assess such a possibility by various investment valuation methods of the project.*

Keywords

Smart Grid, Electricity storage systems, renewable energy sources, investment appraisal of projects.

1. Introduction

The main goal of "Smart Grid" technologies (Smart Grid) create the production, transmission of energy safer and more energy efficient. Smart grids combine modern energy capabilities with innovations in digital technology to manage all energy. Thanks to accurate calculations, losses in the electric grid are reduced and uninterrupted supply is ensured. A distinctive feature of the "Smart Grid" is self-diagnosis and self-healing, thanks to which the system can identify emergency sections of the network and automatically rebuild. One of the important elements of smart grids is electricity storage systems. Electricity storage systems (ESS) are developing a class of high-tech devices that open new opportunities for developing the electric power industry. They are sufficient technologies that create all the conditions for smart energy management [1].

Leading countries are pursuing existing results in introducing new technologies. In the USA (California), it is planned to introduce 1325 MW of storage capacity by 2020. In 2016, National Grid (United Kingdom) purchased a 201 MW power supply system. Regarding the accumulated energy in China, up to 2021, it is planned to introduce 46 GW of ESS. In the United States and China, large-scale production of storage devices is being developed, focused on the saturation of the domestic market, and one massive export deliveries. Navigant Research, the global market for energy storage systems, will reach $80 billion by 2025. In the last 10 years, the volume of the global market of ESS has grown almost three times [2].

2. The use of power storage systems and their types

Such a diverse set of functions and applications of ESS is determined by wide possibilities in the accumulation and subsequent issuance of electricity from milliseconds to hours and days; wide range of power, energy intensity and its ratio; high permissible charge and discharge currents; almost instantaneous switching between different modes of operation; high energy efficiency. These qualities open technological possibilities to implement new power system architectures [3]:

1. Power system with many opportunities for managing schedules of electricity generation through install high-capacity electricity storage systems at network nodes. In such an energy system, ESS play the role of systemic regulating power, which makes it possible to optimally use the generating capacity [4].

2. Internet energy (IoEN), in which active consumers, including those owning ESS and own generation and

having the ability to both, manage consumption and provide the accumulated electricity and power of their ESS, realize free energy exchange and other energy transactions and services on decentralized markets. The following features are characteristic of Internet energy [4]:

1) Active consumers are needed flexibly controlled consumption, having a significant share in the energy balance.

2) Flexible load management of generating and network capacities; the possibility of local quality control of electricity on the side of consumers and groups of consumers;

3) Self-balancing of power in small power systems, including the generation and load operation in non-synchronous mode. Here, all generators and consumers are connected to each other through an energy hub - a device controlling and regulating the primary power balance

4) Efficient integration of generation based on renewable energy, increasing the efficiency and share of renewable energy in the energy balance of the power system;

The using of power storage systems connected to grids and participation in energy markets face current uncertainties in the regulatory framework. In particular, there is no concept of a "system of accumulation of electric power", which regulate relations between subjects in the electric power industry; there is no consideration of the features of technological connection of ESS to networks, the order of their participation in the electricity, capacity and related markets. There is an uncertainty in explain expenses for the purchase of electricity and revenues from the sale of stored electricity if the Grid Company, and approximately aggregation of the systems and take part in their units in the electricity and capacity markets, own the ESS. This limits and slows developing the industry [5].

3. Methods for evaluating investment projects

Evaluation of the effectiveness of investments is given in the form of a conclusion based on the analysis of performance indicators. There are several methods for assessing the attractiveness of investment projects, and therefore several key indicators, which are a certain set of indicators. Each method is based on the same principle - because of the project implementation, the company must make a profit, while various indicators provide an opportunity to characterize the investment project from all sides and meet the interests of various groups of people involved in the investment. Two groups of assessment methods are usually used, with the help of which the indicators listed are determined [6].

3.1 Static methods of evaluation.

They are based on the “Typical Method for Determining the Economic Efficiency of Capital Investments”. It has not lost its relevance today, since the methods are simple in execution and at the same time provide an opportunity to get a clear picture of the effectiveness of the investment, especially in the first stage of the evaluation work [7].

1. A payback period of investment (Payback period, PP).
2. The coefficient of investment efficiency (Account rate of return, ARR).

Static methods of investment evaluation do not go without flaws. The main one is that they do not take into account the time factor, and incomparable values are taken for calculations - the amount of investment in the current value and the profit value.

3.2 Dynamic methods

This group of methods is characterized by complexity and the need to take into account many aspects. They are used to test investment projects of great duration, requiring additional investments in the course of their implementation. When using dynamic methods, an important component is a search for discount rates, which allow you to bring income and expenses to values close to real [7].

3.2.1 Net present value (Net present value, NPV).

A positive NPV is a criterion for accepting an investment project. If it is necessary to make a choice from several projects, preference is given to a project with a larger NPV value.

3.2.2 Profitability index (PI).

Under this indicator, is seen the ratio of the current value of the cash inflow to the net present value of the cash outflow, taking into account the initial investment. If, when calculating, the value of PI is greater than one, then the project should be accepted; if less, it is rejected.

3.2.3 The internal rate of return (IRR)

The internal rate of return (IRR), or internal investment rate of return, is the value of the discount rate at which the project's NPV is zero. For example, if a project is financed by a loan from a commercial bank, then IRR shows the upper limit of the level of bank interest rates.

3.2.4 Discount rate

When choosing a discount rate, you need to take into account inflation, the cost of all sources of funds for

investment and indicators of risk. The method for calculating risk premiums for various classes of investment/investment projects was presented to scientists J. Honko. These risk premiums are presented in an aggregated form, and the investor needs to choose the purpose of the investment and, the risk change [8].

Forced investment	-
Maintaining a market position	1%
Upgrade of the company's fixed assets	7%
cost savings	10%
Development of new projects	15%
Innovative projects	20%

Tab. 1. Discount rate options [8]

4. Creating a calculation model

To analyze the effect of accumulation systems, a village electric grid model was created. This model includes the elements shown in Table 1. The load is expressed in the annual load schedule Fig 1.

Elements in the model	Number of elements	Installed power, kW
Line	91	-
Loads	58	-
battery	4	48
Photovoltaic	58	191,4
Transformer	2	-

Tab. 2. Elements of the model

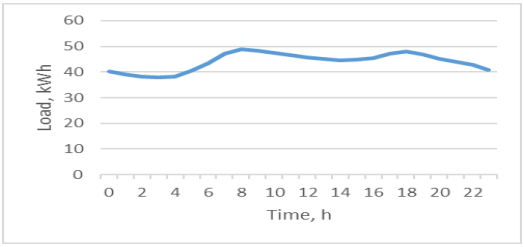


Fig. 1. Annual load schedule

To analyze the investment valuation, prices were taken from one of the leading electricity exchanges in Europe, Nord Pool [9].

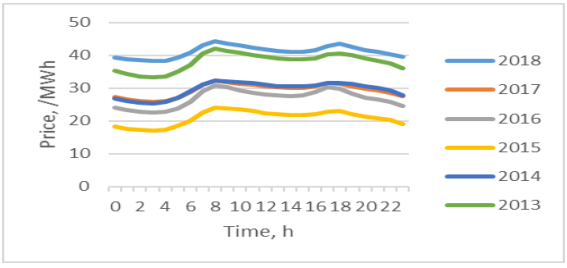


Fig. 2. The annual schedule of electricity prices for the period 2013-2018

5. Analysis of implementing of power storage systems

Because of state support, and a reduction in the cost of solar panels, many experts predict a broader implementation among home consumers. As part of these forecasts, a simple analysis of the power losses in the electrical network of the model under study was carried out. Because of the analysis, dependences of the increase in active and reactive power losses because of the increase in installed power were obtained [10].

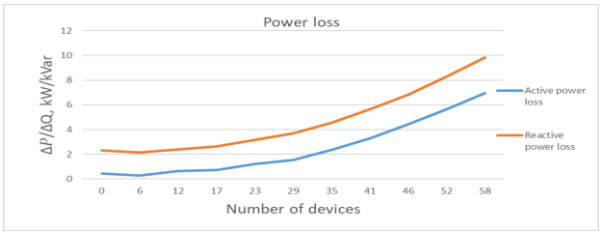


Fig. 3. Power loss versus the number of Photovoltaic

As we see from graph 1, with an increase in the number of solar panels, power losses increase because of the change in the power flow direction. One solution to this problem for power companies is to place energy storage systems. Within this model, the energy storage systems of 48 kW were installed [11].

Mode of Grid	ΔP, kW	ΔQ, kVar
Without elements of a Smart Grid	0,45	2,33
Photovoltaic	6,94	9,85
Battery charging mode	2,61	4,43
Battery recharging mode	0,39	2,18
battery charging mode with Photovoltaic	4,7	7,38

Tab. 3. Power loss in different modes

As we see from Table 2, as a result installing the batteries, it was possible to reduce the power loss during operation of the solar panels by 32.28%, and during the discharge of the battery. However, power losses also increased in a time when the batteries are being charged. The next step in analyzing the benefits of installing energy storage systems is to equalize the daily load curve. For analysis, the day was taken in January and the day in July. Load graphs are presented below in Figure 1 and 2.

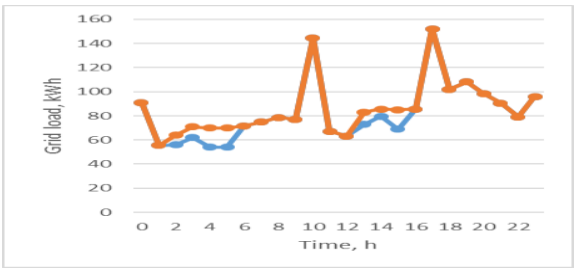


Fig. 4. Daily load chart (January) without using the energy storage system and using the system

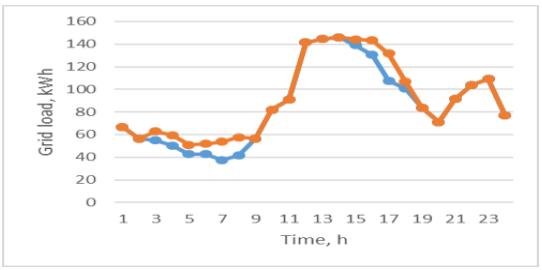


Fig. 5. Daily load chart (July) without using the energy storage system and using the system

As you can see in the picture, the batteries are charged during periods when the load on the network decreases. Thus, the schedule becomes evener. Also not shown on the graph, but at peak times the battery will operate in a discharge mode, which will reduce power loss at the busiest moment, and also reduce the cost of electrical stations. The result may vary depending on the specific day [12].

After Installing the accumulation systems, the power company or consumers will have the opportunity to use price arbitrage. Arbitrage (from Fr. Arbitrage is a fair decision) in the economy - several related transactions aimed at extracting profit from the difference in prices for the same or related assets at the same time in different markets (spatial arbitrage) or on the same volume the same market at different points in time (temporary arbitration, ordinary exchange speculation) [13].

When analyzing prices, it was found that most often the lowest price for electricity is at night and in the afternoon. It is also necessary to take into account that when installing a large number of solar panels, there may be situations when the price of electricity will be much lower than the average price calculated per year. Below are graphs that show the profit from this arbitration analyzed for 5 years. The battery was charged from 2 to 4 and 13-15 and discharged from 8 to 10 and 16 to 18 hours. It is important to note that there may be days when this method will not work. However, on average, over a long period, arbitration remains profitable [14]. The cost of the accumulation system is 10,000 €.

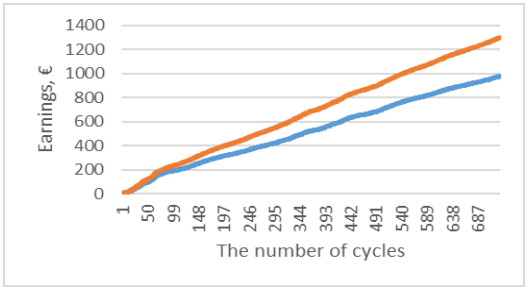


Fig. 6. Profit from the temporary arbitration of electricity prices.

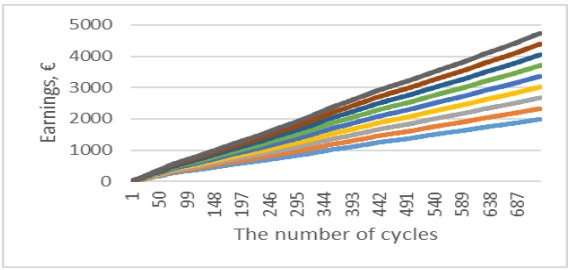


Fig. 7. Profit from the temporary arbitration of electricity prices. Different curves show different possible price reductions because of the development of renewable energy sources.

The graph shows the cases when the price of electricity was about 0 because of the strong production of energy from renewable sources of energy. It can be assumed that these cases will occur more often because of increased efficiency and the number of power plants using renewable energy sources [15].

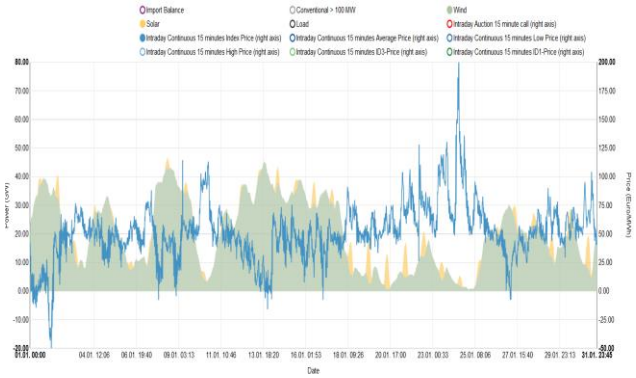


Fig. 8. Price changes in Germany because of high generation in renewable energy sources

It is also worth noting that if a price drops because of strong production on renewable energy sources, the profit from arbitration may increase by a factor of 2-3.

Price changes for photovoltaic		NPV, €	PI	ARR	IRR
1	Normal mode	-8224,8	0,10	0,26	-19%
1	Normal mode+ the power loss	-8224,3	0,10	0,26	-19%
0	Normal mode+PV	-3804,6	0,35	0,88	-2%
	NM+PV+PL	-3804,1	0,35	0,88	-2%
0,2	Normal mode+PV	-4200,6	0,33	0,83	-3%
	NM+PV+PL	-4200	0,33	0,83	-3%
0,4	Normal mode+PV	-4596,5	0,31	0,78	-4%
	NM+PV+PL	-4596	0,31	0,78	-4%
0,6	Normal mode+PV	-4992,4	0,29	0,73	-5%
	NM+PV+PL	-4991,9	0,29	0,73	-5%
0,8	Normal mode+PV	-5388,4	0,27	0,68	-6%
	NM+PV+PL	-5387,9	0,27	0,68	-6%

Tab. 4. The results of the investment assessment methods of the project

It is possible to assess the feasibility of using temporary arbitration of energy storage systems; for this, 4

methods for assessing the investment-attractiveness of the project were chosen. Almost all methods have shown that this investment project is not profitable. Most came in terms of profitability in the situation if in the daily cycle of charging renewable energy sources will produce a large amount of energy, which in theory could reduce the price of electricity to 0€ [16].

6. Conclusion

Now, the widespread use of high-capacity energy storage systems has no economic justification. The following factors prevent widespread use:

1) High cost
2) Inconsistent production of electricity from renewable energy sources
3) The low number of charge/discharge cycles

However, accumulation systems improve network operation parameters, improve power quality, and make the network more flexible to manage. Therefore, perhaps installing accumulation systems makes sense from a technical point of view.

Acknowledgements

This work was supported by the Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences by the projects VEGA No. 1/0372/18.

References

[1] M. SHAMSHIRI, C. K. GAN, A C. W. TAN, "A review of recent development in smart grid and micro-grid laboratories", v 2012 *IEEE International Power Engineering and Optimization Conference Melaka*, Malaysia, 2012, s. 367–372.

[2] U. Udalcova and D. Kholkina. EXPERT AND ANALYTICAL REPORT "THE MARKET OF ELECTRICITY ACCUMULATION SYSTEMS IN RUSSIA: DEVELOPMENT POTENTIAL" [Online]. Available: http://www.rusnano.com/upload/images/sitefiles/files/Condenses_system_markets_in_Russia.pdf

[3] NAISH C, MCCUBBIN I, EDBERG O, HARFOOT M. Outlook of Energy Storage Technologies. Brussels: European Parliament, Policy Department, *Economic and Scientific Policy*; 2008.

[4] SAMI BS. An Intelligent Power Management Investigation for Stand-Alone Hybrid System Using Short-Time Energy Storage. *International Journal of Power Electronics and Drive Systems.* 2017 ;8(1):pp. 367.

[5] CHEN H, CONG TN, YANG W, TAN C, LI Y, DING Y. Progress in electrical energy storage system: A critical review. *Progress in Natural Science. 2009*; 19(3): pp. 291-312.

[6] LILJEBLOM E. VAIHEKOSKI M., "Investment Evaluation Methods and Required Rate of Return in Finnish Publicly Listed Companies," p. 16.

[7] HÄRUS, N. C. "Analyzing energy efficiency investments in the process industry - Case Sachtleben Pigments Oy," p. 120.

[8] HONKO, J. & VIRTANEN, K., 1975, The investment process in Finnish industrial enterprises. A study of capital investment planning and control process in the fifth largest Finnish industrial enterprises. Helsinki School of Economics, Series A-16.

[9] Market data "See what Nord Pool can offer you." [Online]. Available: http://www.nordpoolspot.com/.

[10] AL-SALAYMEH A, AL-HAMAMRE Z, SHARAF F, ABDELKADER MR. Technical and economic assessment of the utilization of photovoltaic systems in residential buildings: The case of Jordan. *Energy conversion and management. 2010*; 51(8): pp. 1719-26.

[11] VARKANI AK, DARAEEPOUR A, MONSEF H. A new self-scheduling strategy for integrated operation of wind and pumped-storage power plants in power markets. *Appl Energ* 2011;88:5002-12.

[12] CONNOLLY D, LUND H, FINN P, MATHIESEN BV, LEAHY M. Practical operation strategies for pumped hydroelectric energy storage (PHES) utilizing electricity price arbitrage. *Energ Policy 2011*;39:4189-96.

[13] IBRAHIM H, BEGUENANE R, MERABET A. Technical and financial benefits of electrical energy storage. *IEEE Electrical Power and Energy Conference (EPEC)*, 2012; pp. 86-91

[14] HELLSTRÖM J, LUNDGREN J, YU H. Why do electricity prices jump? Empirical evidence from the Nordic electricity market. Energy Econ 2012;34:1774-81.

[15] Market data, Fraunhofer Institute for Solar Energy Systems ISE [Online]. Available: https://www.energy-charts.de/price.htm

[16] WALAWALKAR R, APT J, MANCINI R. Economics of electric energy storage for energy arbitrage and regulation in New York. *Energ Policy 2007*;35:2558-68.

About Authors

Maksym Oliinyk was born in 1994. In 2018 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in electric power engineering on subject Smart Grid. At present is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. His scientific research is mainly focused on design and researching SMART GRIDS.

Daniel Pál was born in 1994. In 2018 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in electric power engineering on subject design for the reconstruction of school spaces lighting. At present is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. His scientific research is mainly focused on research of renewable resources of energy and for the reduction of the power losses in SMART GRIDS.

Model of the microgrid for the future smart grid

Michal IVANČÁK[1], *Juliana IVANČÁKOVÁ*[2]

[1] Dept. of Electric Power Engineering, FEI, Technical university of Košice, Letná 9, 042 00 Košice, Slovak Republic
[2] Dept. of Cybernetics and Artificial Intelligence, FEI, Technical university of Košice, Letná 9, 042 00 Košice, Slovak Republic

michal.ivancak@tuke.sk, juliana.ivancakova@tuke.sk

Abstract. *Smart grid or intelligent networks are currently the most often mentioned theme, according to this is research in this area expected. In this paper is describing model of the electrical network through graphical software that replaces mathematical computation, facilitates and speeds up this process. The data for this project are used from real distribution system from the concrete village in Slovakia. In this model is tested influence of renewable energy sources to the distribution grid. On this basic can be build smart grid of the future.*

Keywords

microgrid, renewable energy source, smart grid, smart grid projects.

1. Introduction

In these years "Smart" is as a very often mentioned term. It usually refers to special products with an increasing degree of computing power and intelligence. We know smart products such as smart phone, smart TV or other smart electronics with a variety of features and applications. Electrical energy is currently a strategic "raw material" and its importance in the future years is clearly growing. For this reason, increasing emphasis is placed on the stability, safety and security of the electricity supply to end customers. Therefore, the computerization is increasingly being introduced into the electricity system during the time. Very popular is also the new term “Smart Grid”.

The many of projects, articles and publications focus on theme of Smart Grid, as well as foreign and domestic conferences. Despite the great popularity, there is a wide inconsistency in the definition of this term. Nevertheless, Smart Grid is often referred also to as the network capable of using more renewable energy sources and distributed production than the current network.

Comprehension of the current Smart Grid network is rather difficult and from costs perspective expensive. It is a long-term process that binds capital over many years. Therefore, it requires a strong commitment from all stakeholders. In addition, it is still not fully verified how the individual technologies within Smart Grid will work together.

2. Implementation of the smart grid

Smart Grid networks have the following features and benefits over classic networks. The biggest difference is the different network topology due to the inclusion of distributed production that causes different energy flow directions. The change is also thanks to the new technologies, two-way communication and the presence of active elements and sensors throughout the system, self-monitoring as well as rapid detection and localization of failures.

Thanks to new technologies, higher reliability, better security, greater convenience for customers and higher efficiency in the use of electricity are expected. Intelligent systems also envisage semi-automatic renewal and auto-regeneration as well as adaptive protection and isolation of a potential problem. Customers are thus provided with the integration and provision of new services. From the distribution point of view, it is the use of centralized resources along with decentralized resources.

Decentralized sources of small capacities deployed across Europe are in line with the European Union's commitment and commitment to increase its share of renewable energy production to 20% by 2020. There is expected reducing system losses and the associated increase in ecology, economy and operational efficiency as well as support for scattered production along with the development and research of new management methods. results.

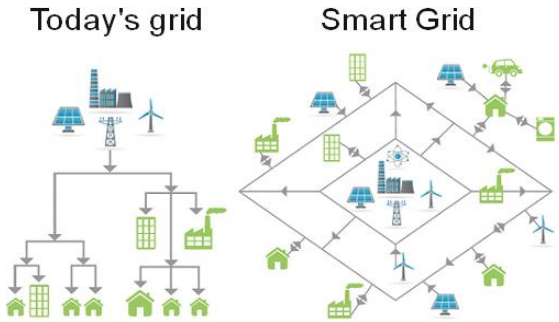


Fig. 1. Comparation of today’s grid and the smart grid.

3. Modelling microgrid

Modelling and using of modelling software are a very good method for making designs thanks to computer calculating performance, especially taking in account complexity of the environment. Of course, in the background of every software there is a set of mathematical formulas that we can use to make modifications more efficiently while changing a constant. So, it is not necessary to repeat the entire calculation.

The necessary part of creating an electrical network is adequate preparation which is needed to ensure that the network's functionality is properly verified to avoid unnecessary investment costs. Without suitable training it would not even be possible. At the same time, the software tools create the right conditions for laboratory testing as well as back-up verification of the functionality of the already implemented project. The basis of the modelling must be precise as possible in order to ensure the accuracy of the results is also close as possible to the reality.

Simulink is a MATLAB extension for simulation and modelling of dynamic systems. It provides the user with the ability to quickly and easily create dynamic system models in the form of block diagrams. Models can be described by equations or can be assembled from blocks representing real system elements. Besides models of physical systems, it is possible to model also control system algorithms including their automatic tuning, signal processing systems, communication and image processing.

The microgrid is divided into four important parts: A diesel generator (15MW), acting as the base power generator; A PV farm (8 MW) combined with a wind farm (4,5 MW), to produce renewable energy; a vehicle to grid system installed next to the last part of the topology which is the load (10 MW) of the grid. The size of the microgrid represents approximately a community of a thousand households during a low consumption day in spring or fall. There are 100 electric vehicles in the base model which means that there is a 1:10 ratio between the cars and the households. This is a possible scenario in a foreseeable future.

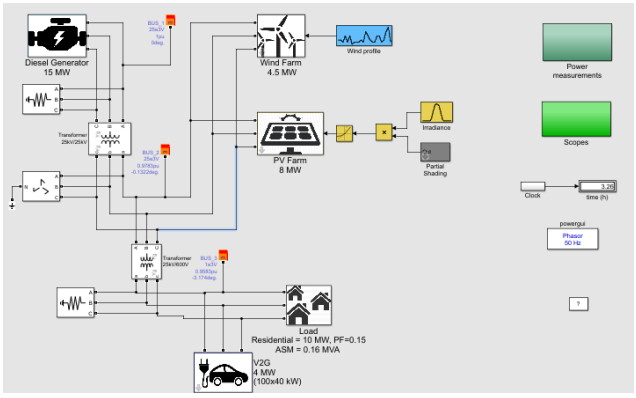


Fig. 2. Model of micro grid in Matlab - Simulink

The diesel generator balances the power consumed and the power produced. We can determine the frequency deviation of the grid by looking at the rotor speed of its synchronous machine.

The load consists of a residential load and an asynchronous machine that is used to express the impact of an industrial inductive load (such as a ventilation system) on a microgrid. Residential load monitors the consumption profile with a given power factor. The asynchronous machine is controlled by a square relationship between rotor speed and mechanical torque.

The photovoltaic power plant produces three energy factors: the size of the area covered by the photovoltaic power plant, the efficiency of solar panels and irradiation data. The simplified model of a wind farm produces electricity after a linear relationship with the wind. When the wind reaches the nominal value, the wind farm produces nominal power. The wind power plant emerges from the grid when the wind speed exceeds the maximum wind speed until the wind returns to its nominal value.

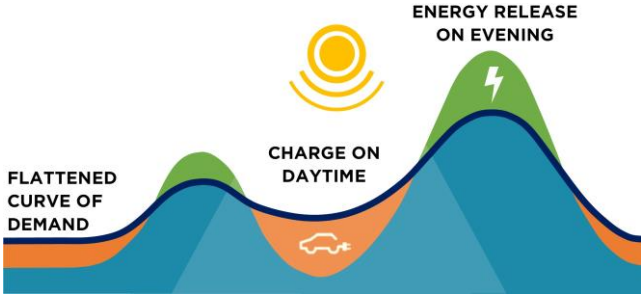


Fig. 3. Optimal plan of charging EV [15]

The V2G has two functions: Controls the charge of the batteries connected to it and uses the available power to regulate the grid when an event occurs during the day. With an increased number of electric vehicles, there will be a situation that can significantly affect the daily load diagram of the power grid. From this point of view, there are several EV charging scenarios. These scenarios consider target variables such as the car owner's working time as well as whether they have the option to recharge their e-car at work or not. The block implements five different car-user profiles:

Profile #1: People going to work with a possibility to charge their car at work.

Profile #2: People going to work with no possibility to charge their car at work.

Profile #3: People going to work with a possibility to charge their car at work but with a longer ride

Profile #4: People staying at home.

Profile #5: People working on a night shift.

The diagram below shows the graphs of charging the car. The graphs are only approximate, but better illustrate the flow of energy in car batteries.

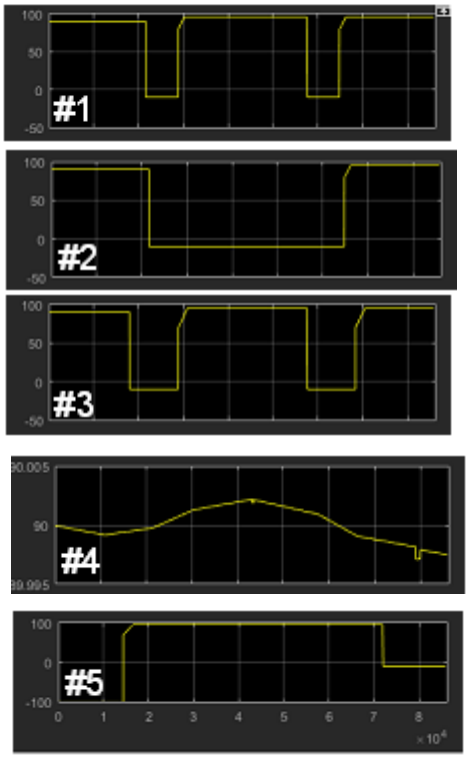


Fig. 4. Comparison of the apparent and active power of a photovoltaic power plant, a wind power plant and a diesel generator.

Run the model and observe the various range signals inside of it. It is possible to monitor rotor speed behaviour in the range above the model.

Click on the Scopes and Power measurements subsystem to access information from different nodes. The charging status of each vehicle profile is also available in this subsystem. Negative charge state means the car is on the road or not connected.

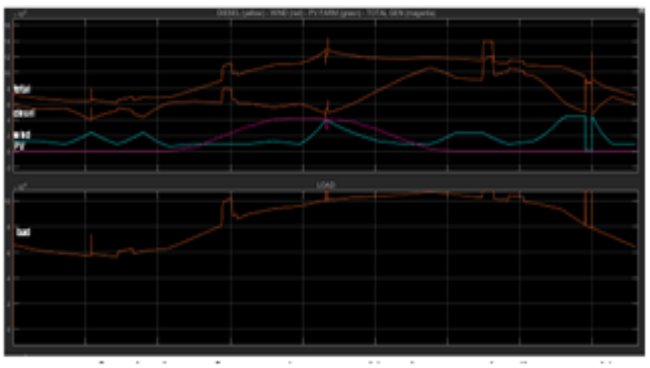


Fig. 5. Processes of production performance (upper graph) and consumption (lower graph)

Simulation takes 24 hours. The intensity of sunlight is controlled by the normal distribution where the highest intensity is reached at noon. In figure it has magenta colour. The wind (blue) changes significantly during the day and has several peaks and minima. Diesel generator (red) holds frequency and naturally regulates electricity generation. Residential load has a typical formula similar to normal household consumption. It is low during the day, increases to the peak during the evening and decreases slowly at night. Three events affect network frequency during the day:

-the start of the asynchronous machine in the third hour

-a partial cloudiness at midday affecting the production of solar energy

-a wind farms cut off in 22 hours when the wind exceeds the permitted maximum permitted wind energy

4. Conclusion

This paper describes a simulation that includes power sources like diesel generator, a photovoltaic and wind power plant, electricity consumption and a model of electric vehicle charging as reserve electricity in case of network drops or surpluses. The system also includes unpredictable sources that in combination with the diesel generator and the electric car system keep the network running.

Because of the model is the off-grid system it is not connected to a larger system, it depends on the reliable operation of the largest source. The base is a diesel generator that is not dependent on wind and solar power but provides the maximum space for energy from renewable energy sources. Several measurements have been performed within the model, demonstrating the functionality of the model and its stability under the given conditions. On a given model, it would still be appropriate to monitor the quality of electrical energy, especially the frequency, since large frequency fluctuations have an undue influence on the functionality of the elements in the network.

The theme of microsite and smart grid modeling is highly up to date due to the speed of this time. Electricity plays a significant role in society, and a small one is aware of its daily needs. Model testing improves network reliability and reduces downtime.

Acknowledgements

Research described in the paper was supervised by Doc. J. Kurimský, FEI TUKE in Košice and supported by the Ministry of Education, Science, Research and Sport of the Slovak Republic and The Slovak Research and Development Agency under the contract No. APVV-17-0372.

References

[1] SIOSHANSI F. P. et al., *Smart Grid: Integrating Renewable, Distributed & Efficient Energy*, 510 (Academic Press, 2011)

[2] VOLČKO V., Smart Grid: Vplyv na prevádzku, bezpečnosť a stabilitu elektrizačnej sústavy, (2015), <http://www.fei.stuba.sk/docs//2015/autoreferaty/Volcko_autoref.pdf >

[3] BORLASE S., *Smart Grids: Infrastructure, Technology and Solutions*, 577, (CRC, 2013)

[4] MOMOH J., *Smart Grid: Fundamentals of Design and Analysis*, 216, (IEEE P., 2012)

[5] M. Shamshiri, C. K. Gan, a C. W. Tan, "A review of recent development in smart grid and micro-grid laboratories", v 2012 *IEEE International Power Engineering and Optimization Conference Melaka, Malaysia*, 2012, s. 367–372.

[6] N. Nikmehr a S. N. Ravadanegh, "Optimal Power Dispatch of Multi-Microgrids at Future Smart Distribution Grids", *IEEE Transactions on Smart Grid*, roč. 6, č. 4, s. 1648–1657, júl. 2015.

[7] K. Zhou, S. Yang, a C. Shen, "A review of electric load classification in smart grid environment", *Renewable and Sustainable Energy Reviews*, roč. 24, s. 103–110, aug. 2013.

[8] G. Lobaccaro, S. Carlucci, a E. Löfström, "A Review of Systems and Technologies for Smart Homes and Smart Grids", *Energies*, roč. 9, č. 5, s. 348, máj. 2016.

[9] N. Nikmehr a S. N. Ravadanegh, "Optimal Power Dispatch of Multi-Microgrids at Future Smart Distribution Grids", *IEEE Transactions on Smart Grid*, roč. 6, č. 4, s. 1648–1657, júl. 2015.

[10] K. Moslehi a R. Kumar, "A Reliability Perspective of the Smart Grid", IEEE Transactions on Smart Grid, roč. 1, č. 1, s. 57–64, jún. 2010.

[11] H. Yang, J. Zhang, J. Qiu, S. Zhang, M. Lai, a Z. Y. Dong, "A Practical Pricing Approach to Smart Grid Demand Response Based on Load Classification", IEEE Transactions on Smart Grid, roč. 9, č. 1, s. 179–190, jan. 2018.

[12] A. Molderink, V. Bakker, M. G. C. Bosman, J. L. Hurink, a G. J. M. Smit, "Management and Control of Domestic Smart Grid Technology", IEEE Transactions on Smart Grid, roč. 1, č. 2, s. 109–119, sep. 2010.

[13] M. Hashmi, S. Hänninen, a K. Mäki, "Survey of smart grid concepts, architectures, and technological demonstrations worldwide", v 2011 IEEE PES CONFERENCE ON INNOVATIVE SMART GRID TECHNOLOGIES LATIN AMERICA (ISGT LA), 2011, s. 1–7.

[14] Mathworks, Documentation of using Matlab, https://www.mathworks.com/

[15] The solution to sustainable urban mobility and energy; Office AEL. Online: http://www.amsterdamvehicle2grid.nl/

About Authors...

Michal IVANČÁK was born in 1993. In 2016 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. At present he is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master's degree in electric power engineering on subject of the selectivity analysis for electrical balance of power plant with coordination with ABB. His scientific research is mainly focused on research of smart grid.

Juliana IVANČÁKOVÁ was born in 1993. In 2017 graduated (MSc) at the Department of Cybernetics and Artificial Intelligence on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. At present she is a Ph.D. student in the Department of Cybernetics and Artificial Intelligence on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. She received a master's degree in Business Informatics on subject of Cancer diagnostics through suitable data mining methods. Her scientific research is mainly focused on research of models and methods of data analysis for extraction of semantic representation of data sources.

Simulation of daily, artificial and combined lighting in the software DIALux

Daniel PÁL[1], *Jakub URBANSKÝ*[2], *Maksym OLIINYK*[3]

[1,2,3] Department of Electrical Power Engineering, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Mäsiarska 74, 041 20 Košice, Slovak Republic

daniel.pal@tuke.sk, jakub.urbansky@tuke.sk, maksym.oliinyk@tuke.sk

Abstract. *This article deals with the simulation of daily, artificial and combined lighting in a room. Simulation is performed within the software DIALux. With usage of the simulation software DIALux it is possible to simulate the intensity of light changes during the day, month and year. It can also be used to simulate combined lighting. With performed simulation it's possible to estimate economical and energy savings of combined lightning. Each simulation should be performed in a way that provides the minimum luminous intensity according to standard EN 12464-1. After evaluating the simulated data, these data can be used as a reference to calculate the savings when combined lighting is realized in real conditions. The main goal of this article is comparison between uncontrolled and controlled artificial lighting shown on an example of one room. The results are showing possible economical and energy savings that would be incurred with case of using the combined light instead of purely artificial light.*

Keywords

DIALux, artificial lighting, energy saving, sensor.

1. Introduction

Lighting plays an important role in human live. When a person is outside, in nature or when he is in a room, the light is always present. The only element constantly changing is the type of lightning used. During office work, the most common type of lighting is artificial lighting. We are using luminaires to create artificial lighting. It is very important that to have sufficient amount of indoor lighting, to perform out daily tasks effectively. If the lighting is not sufficient, e.g. there is not enough illumination in the room, the worker will be tired which also affects his work performance. Other factors that are decreasing work performance are for example the incorrect placement of a desk, an unsuitable office chair, insufficient hydration, lack of fresh air in the room or insufficient ambient temperature. The article will be dealing with artificial lighting of one room and there will be carried out of simulation of possible energy consumption and economical savings with deployment of combined lighting instead of solely artificial one. Reducing electricity consumption will not affect the minimum light intensity that needs to be achieved according to EN 12 464-1.

2. Lighting types

Lighting can generally be divided into 3 types:

- daily,
- artificial,
- combined.

2.1 Daily lighting

Daylight is the lighting that comes directly from the sun. The sun is a source of daylight, so it can also be called natural lighting. In general, it can be divided into 3 parts, depending on how the light comes into the room.

- *Direct lighting* - it comes directly from the sun and enters the room without reflecting.
- *Diffuse illumination* - is lighting that comes from the sun and is reflected from the object. (e.g. table, chair, ...).
- *Reflected illumination* - is lighting that is reflected before entering the room [1] [2].

Daylight intensity is not the same all day; it is constantly changing throughout the day. The greatest intensity values can be measured at around 12:00, the smallest in the morning or in the evening. In the night there is no daylight. Time is not the only parameter, on which the intensity of lighting depends. It depends on several factors:

- *Sky Status* - clear, partly cloudy or overcast.
- *Latitude* - maximum hours per year the sun is shining around the equator and the daytime light is reduced when the equator is moved to the north or to the south.
- *Seasons* - in the winter months are the shortest days, in the summer months are the longest, so the amount of daylight depends

on the season. In the summer months, daylight can be used more efficiently because it is available for a longer time [3].

2.2 Artificial lightning

Artificial lighting is generated by artificial sources, by luminaires. Light from luminaires strikes directly and after reflection from surrounding surfaces to a place, to a room, where we need [2]. Artificial lighting is required for us to work during hours, when daylight is very low, for example in night, in mornings or evenings.

The requirements for artificial lighting in the room are described in the standard EN 12 464-1 and parameters, which are in the standard we needs to be considered, when are designing lighting.

2.3 Combined lighting

Daylight is not sufficient in any case because its value changes during the day. In those periods when there is not enough of daylight, it is necessary to supplement it with artificial lighting and then a combined lighting will be created [3]. The best solution for combined lighting is to use a sensor that will monitor the amount of daylight in the room. Based on sensor data, artificial lighting is controlled to reduce room energy consumption.

3. Simulation software

For carrying out necessary simulation the software DIALux in version 4.13 was. Using the software you can design and visualize the room, and based on this input data calculates the distribution of light intensity in the room. The program is freeware, even though is frequently used in practice [4]. Recently, designers have also started using the program in designing various projects such as designing a room, building, road, public lighting or sport stadiums. The reason for his popularity is that the software is easy to use.

When the proposed project will be built in real life, for example, a stadium will be built or there will be realized reconstruction of the lighting system in the room is necessary to design the project to the smallest details to ensure that the results calculated by the program will correspond to the results in reality. In the program, there are various objects that are predefined. In addition, we have the ability to correct these objects to specify object parameters such as length, width, height or color.

4. The simulation of different types of lighting in the software DIALux

The first step is to design the analyzed room. It is necessary to set the correct room parameters such as:

- width,
- length,
- height.

When these settings are already constructed, the editing of the room is completed. After editing the room, it is necessary to insert windows and doors into the room, and also other objects such as cabinets, tables, chairs, plants. When everything is inserted and set, it is necessary to define the computational surfaces so that the program knows where to calculate the parameters that are important to follow. Floor plan of a room with windows, doors, tables, chairs and with calculate surfaces can be seen in the Fig. 1.

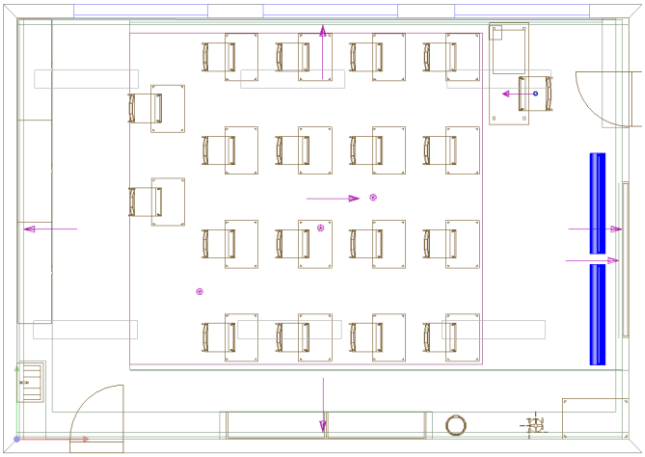


Fig. 1 Floor plan of a room with objects and with calculate surfaces

In the software we can see the room also in 3D format, which we can see in the Fig. 2. This makes it easier to understand the structure of the room.

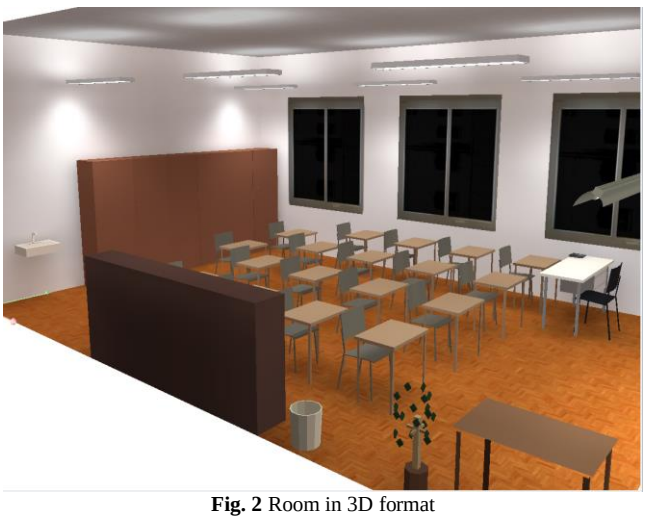


Fig. 2 Room in 3D format

4.1 Daylight simulation

Simulation of daylight in the program is a bit complicated because other parameter needs to be taken in to consideration. This parameter is coordinates of the room. Intensity is depending on specific location. In addition, it is also necessary to set the northern orientation of the building. It is also necessary to insert light scenes so that

the program knows how much time we need to count on the program. After all parameters have been set, the software calculates the lighting illuminance distribution in the room. Fig. 3 is illustrating the amount of daylight based on season, which are represented by their typical months.

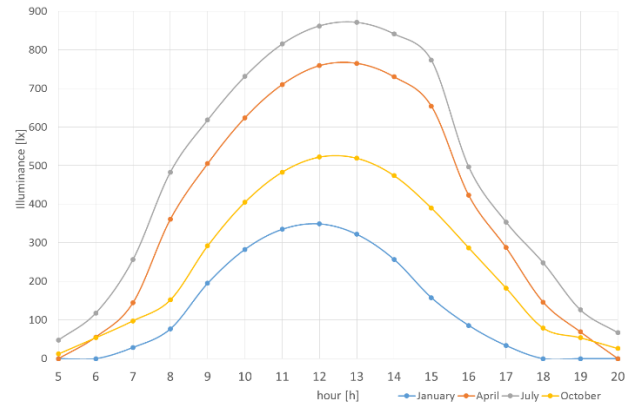


Fig. 3 Dependence of light intensity during the day in different seasons

4.2 Artificial lighting simulation

The simulation of artificial lighting is most important. The room has to be designed to meet legislative requirements. Each building, each room has different requirements which is described in standard EN 12 464-1. The room described in this article is a class, where it is important to reach a minimal value of illuminance at least 300 lx [5]. It is necessary to try different luminaires from different manufacturers, different shapes and nominal power in the simulation. Finally, the lamp from the manufacturer PHILIPS was selected. To reach the minimal value of illuminance, it was necessary to add 6 pieces of lamp to the program and to set it up to work at 100 % power.

4.3 Combined lighting simulation

The brand new room are applied and sensors that monitor the amount of daylight in the room and on the basis of measured data adjustment of the power lamps. Due to the fact that the lights do not shine at 100% power, it is possible to save part of the electricity that is used for lighting. For the simulation combined lighting is required to simulate daylight and artificial light. In the case of artificial lighting, it is also necessary to simulate the percentages of luminaires from 0 % to 100 %. This is necessary in order to calculate what energy savings would be. The combination of the results was made with the following procedures:

1. It is necessary to see in the standard the intensity of lighting we should have in the room.
2. It is necessary to evaluate the results of light intensity when daylight.
3. Will be calculate the difference between the value in the standard and achieved result.
4. Missing intensity of lighting must be illuminated with the lamps.

As an example it is possible to mention. If the daylight intensity of the room is 100 lx. The normative requirement is requiring at least 300 lx. It means that 200 lx needs to be illuminated artificially. For instance, at 70% of nominal power can luminaire achieve additional 200 lx which is required with normative. With such regulation there would be achieved 30% energy saving in comparison with case without using of regulation.

With this procedure, all the days and all the times of the year need to be done to determine what savings would be made in the year with regulation. In real life, there are various sensors from different manufacturers that monitor daylight in the room. It is important to choose the right one because not all sensors work with all of luminaires. Since in the simulation it is used PHILIPS luminaries, so it would be appropriate to use sensors their sensor. One option is to use PHILIPS LRL1220 sensors. The price of one sensor is approximate 20 € [6].

5. Comparation of operating costs without and with sensor

In order to assess that the use of sensors is advantageous or not, the results need to be compared to operating costs. Operating costs determine how much electricity will be used to illuminate a room. Operating costs can be calculated with Equation (1):

$$OC = TLO * t * d * ep \quad (1)$$

Where:

OC - operating costs [€].
TLO - total light output of the room [kW].
t - time of day of the room lighting (t = 8 hours / day).
d - number of days. It was calculate a 190 days.
ep - electricity price [€ / kWh].

Electricity consumption vary from month to month, as shown in Fig. 4. The reason for these changes is that the amount of daylight in the room is changing. Most of the daylight is in the summer months. The least is in the winter months. This can also be seen in Fig. 4, as the biggest savings are in achieved during the months May and June, the smallest in January. However, electricity consumption also depends on how many days are workdays during month.

Financial savings are increasing with electricity price. Electricity price will increase in future; thus economical savings are expected to be greater. Therefore, is necessary to emphasize the importance of lighting control. This is shown in Fig. 5.

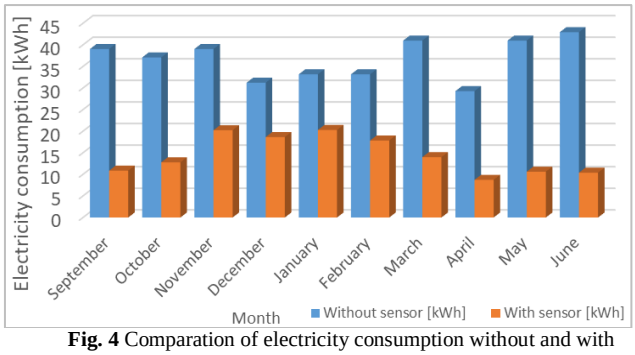


Fig. 4 Comparation of electricity consumption without and with sensor

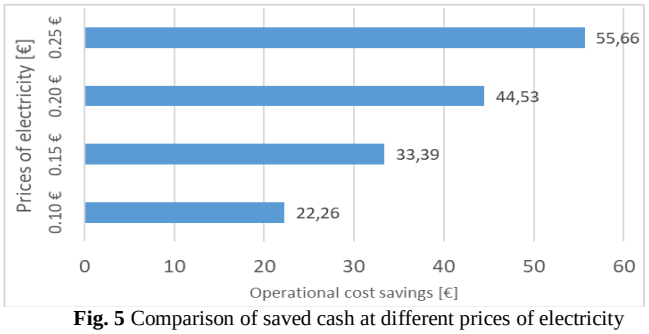


Fig. 5 Comparison of saved cash at different prices of electricity

6. Conclusion

This article deals with the simulation of daily, artificial and combined lighting in a room. It has been explained how to simulate the different types of lighting in the software Dialux. It has also dealt with pooled lighting, and it has also been shown how the intensity of lighting changes over the course of the day and what consequences it has on reducing operating costs. Information about saving energy and thus money is theoretical. The program room was made as accurately as possible in order to achieve approximately the same results if the proposal is actually implemented. The room, which was simulated in the software DIALux has a standard area of 30 m^2. It has 3 windows, from where will come the daylight. In 1 year the operational saving costs will be a 45 €, when price of electricity 0,20 €/kWh is considered. With increasing price of electricity (e.g. 0,25 - 0,3 €/kWh or greater) operational cost savings increases as well. The amount of saving also depends on how many window and luminaires are present in the room. It can be seen that it is possible to save money with the sensors. If we are seeing a larger period, for example 5 or 10 years, the savings would be greater. The return on investment depends on whether only sensors or even sensors and lights are bought. If only sensors are bought (2 – First and second sensor will control 3 - 3 luminaires) the return of investment is 1 year, because 45 € will be saved and estimated price for 2 sensors is 40 - 50 €. If sensors and luminaires are bought, then return of investment depends on the price per luminaire. Return investment for the sensors in this case 1 year as well, but return investment for the luminaires is a longer than 1 year.

Acknowledgements

Research described in the paper was supervised by doc. Ing. Ľubomír Beňa Phd., KEE FEI TUKE and supported by the Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences by the projects VEGA No. 1/0372/18.

References

[1] Daylight Strategies [online]. Available at: https://www.educate-sustainability.eu/kb/content/daylight-strategies. Accessed on 7. March 2019.

[2] SOKANSKÝ, Karel et al.: Světelná technika. I. vyd. Praha: ČVUT, 2011. 256 s. ISBN 978-80-01-04941-9.

[3] Osvetlenie v priemysle [online]. Available at: https://www.siea.sk/letaky/c-3905/osvetlenie-v-priemysle/. Accessed on 19. March 2019.

[4] DIALux family. For lighting designer and manufacturer [online]. Available at: https://www.dial.de/en/dialux/. Accessed on 18. March 2019.

[5] EN 12464 – 1. Light and lighting.Lighting of work places. Part 1:Indoor work places.

[6] LuxSense LRL 1220 TL5 [online]. Available at: https://www.assets.lighting.philips.com/is/content/PhilipsLighting/fp913700182182-pss-global. Accessed on 19. March 2019.

About Authors

Daniel PÁL was born in 1994. In 2018 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in electric power engineering on subject design for the reconstruction of school spaces lighting. At present is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. His scientific research is mainly focused on research of renewable resources of energy and for the reduction of the power losses in SMART GRIDS.

Jakub URBANSKÝ was born in 1991. In 2017 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. Currently is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. His scientific research is mainly focused on research of renewables sources of energy.

Maksym OLIINYK was born in 1994. In 2018 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in electric power engineering on subject Smart Grid. At present is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. His scientific research is mainly focused on design and researching SMART GRIDS.

Design of a measurement assembly and methodology for determining the conductor's current load

Michal ŠPES[1], *Jakub Urbanský*[1] , *Daniel Pál*[1] , *Michal MÁRTON*[2]

[1] Department of Electrical Power Engineering, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Mäsiarska 74, 041 20 Košice, Slovak Republic
[2] Department of Electronics and Multimedia Telecommunications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Vysokoškolská 4, 040 01 Košice, Slovak Republic

Michal.Spes@tuke.sk, Jakub. Urbansky@tuke.sk, Daniel.Pal@tuke.sk, Michal.Marton@tuke.sk

Abstract.

The constant increase in the number of electricity sources requires increasing transmission capabilities of transmission lines, whether transmission or distribution networks. Member States of the European Union are committed to reducing greenhouse gas emissions and also to making renewable energy 20% of final energy consumption. This leads, on the one hand, to limiting the combustion of fossil fuels, but also to limiting the production of nuclear power plants. In 2011, Germany announced the closure of its nuclear power plants by 2022 at the latest, which, due to the energy infrastructure of Germany, causes problems in the management of power systems in neighboring countries. The second problem is the unpredictable or severely limited prediction of electricity generation from renewable sources. While the location and construction of some of the new power sources is relatively simple, the construction of new power lines is, on the other hand, more demanding in terms of time or money. As a result, we are continually looking for new ways to increase the transmission capabilities of existing lines, taking into account the lowest investment costs and minimizing losses. It is possible to determine the driver's current carrying capacity based on ambient conditions: Using simulation programs and modeling drivers and environmental conditions, measuring on the selected line, Measurement on a measuring set in laboratory conditions. This article deals with the design of a meter assembly and methodology for determining driver current load. In the paper, the measurement assembly and the measurement procedure are described together with the evaluation of the current load measurement for the 24-AL1 / 4-ST1A wire.

Keywords

Power system, Power lines, Maximum current, Overhead lines, Climatic conditions,

1. Introduction

Power system is a set of interconnected facilities used for the generation, transformation and transmission of electricity. Transmission of electricity is provided at the level of the transmission system and level of the distribution system. Within the transmission systems are mostly used twisted ACSR ropes, which, depending on the voltage level can be arranged in a bunch or individually.

Each ACSR rope has determined the maximum allowable current that it can flow without any damage. Maximum allowable current depends on the following parameters:

Electrical and mechanical properties of the conductor material

- Thermal insulation properties of the material in conductor
- The thermal capacity and conductivity of the guide body and its ability to transfer heat to the surroundings
- Environmental conditions

For the purposes of this article, we will examine the impact of solar radiation on the maximum current value.

For the design of power lines, applies standard EN 50341, which for the calculation of the maximum current value determined the following environmental conditions.

- The current conductor is the highest loaded
- The ambient temperature is 35 °C
- Wind speed is 0.5 m / s at 45 ° angle of impact
- Global temperature solar radiation is 1000W/m2
- Absorption coefficient is 0.5
- Emissivity coefficient is 0.5

2. The steady state temperature of the electrical conductor

The maximum permissible current value of the conductors is not constant but varies over time, depending on the ambient conditions.

The maximum permissible current value is defined as the current value that can flow an electrical conductor without damaging or disrupting its function. Disrupting its function is due to exceeding the maximum permissible operating temperature.

The thermal equilibrium of conductor is determined by the following equation, wherein the left part of the formula is formed by factors, which heat the conductor and the right side of the equation is made by factors which cool the conductor.

$$P_Z + P_S + P_C = C_v \cdot \frac{d\Theta}{dt} + P_k + P_r + P_w \qquad (1)$$

Where:

P_J (W/m) - heat losses in the conductor

P_M (W/m) - magnetic heating of magnetic field variations AC

P_S (W/m) - solar radiation

P_i (W/m) - heating from the corona

P_C (W/m) - cooling by heat convection – by radiation

P_r (W/m) - radiant cooling

P_W (W/m) - cooling from water evaporation

From the above expression (1) can be expressed equations for calculating current, while respecting environmental conditions.

$$I = \sqrt{\frac{P_r + P_k - P_S}{R_{ac}}} \qquad (2)$$

In the calculations heating due to the corona neglected. The heating of conductor due to the corona (P_C) occurs mainly during the rain and wind, which is cooling conductor highest. It is also possible to neglect cooled by evaporation of water (P_w).

3. Design of experimental set for the verify mathematical model for determination of current capacity of conductor

The basic hypothesis for this work is the verification of the proposed mathematical model, which was determined on the basis of empirical equations to determine the current carrying capacity of drivers based on the influence of ambient conditions according to the CIGRE standard.

The main element for exploration will be the driver that will run the I-th current value. This current, due to Joula's losses, causes the driver to heat up while neglecting ambient conditions. On the basis of theoretical knowledge, when the influence of ambient conditions on the temperature of the conductor is influenced, the ambient temperature, which also enters the temperature contribution due to Joule's losses, and the intensity of the solar radiation influence. On the other hand, when the driver cools down, the temperature drop due to forced air flow and the temperature drop due to natural convection are affected.

Another essential element for the measuring assembly is a transformer, which allows continuous current regulation up to the current I-value along with the current flow measurement. The measurement will also include thermal probes for measuring conductor temperature on the surface.

On the basis of the theoretical analysis for the measuring assembly, a measuring opinion was compiled. The block diagram of the measurement assembly is shown in the following figure.

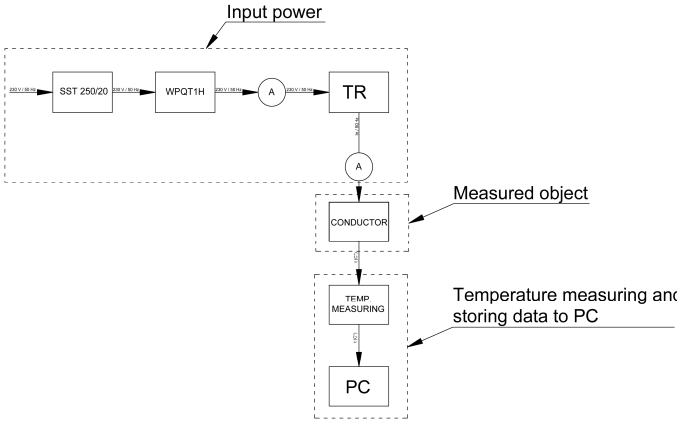


Fig. 1 Block diagram of equipment for determining ampacity

There is a three-phase power supply SST 250/20 with input voltage regulation 0 - 250V and current regulation with range 0 - 20A at the input of the measuring system. The SST250 / 20 power supply is a high current transformer. At the output of the large current transformer, the current is regulated in the range of 0 - 1000A label. WPQT1H is connected in the wiring between the SST250 / 20 drive and the large current transformer. The device primarily serves as a test device for electromechanical protection. In the measuring system, its protective functions are used to prevent damage to the large current transformer.

On the secondary side of the large current transformer, the examined AlFe rope is connected together with temperature probes for temperature measurement. Based on the driver's temperature, the maximum permissible driver temperature can be determined.

A flowchart measurement procedure is shown in Fig. 2. After initializing the ampacity measurement, the temperature and humidity measurement in the room and the driver temperature are started simultaneously. Subsequently, the I-th nominal current value is set by the voltage source depending on the purpose of the measurement. To terminate the measurement, the condition of the driver's temperature is set, ie. has no upward or downward tendency.

4. Determination of the influence of ambient conditions on current carrying capacity of conductors by measurement

An experiment was performed based on the proposed experimental setup and flowchart to measure the influence of ambient temperature on drivers current load. Due to the technical and safety constraints of the laboratory, experiments were carried out on ACSR ropes used in HV grids.From the point of view of ensuring the credibility of the measurement results, individual measurements were made with a 10-fold repeat.

The surrounding conditions were as follows:

- Ambient temperature Ta = 27°C
- Wind speed v = 0 $m.s^{-1}$,
- Sunlight intensity ls = 0 $W.m^{-2}$.

The measured temperature was arithmetically averaged.

Parameters of examined ACSR conductor are in table 1.

Results from individual measurements are at Fig. 3

In the verification of the mathematical model, measurements were made on the basis of the proposed experimental set-up and technological procedure for measuring the power of the external lines.

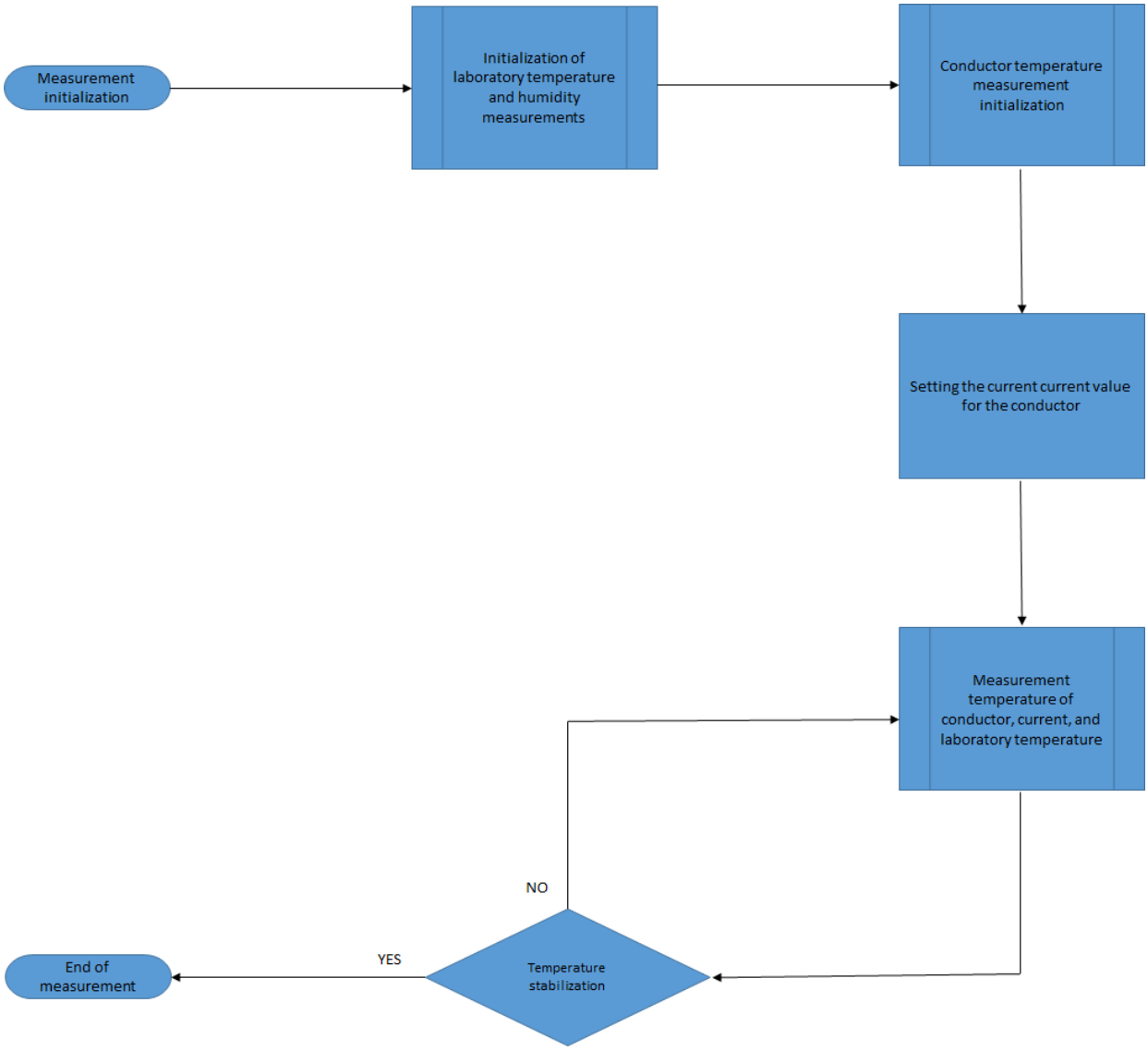


Fig. 2 Flowchart for power ampacity research

Examined conductor		24-AL1/4-ST1A
Crossection (mm^2)	AL	23,64
	ST	3,94
	Overall	27,58
Number of wire	AL	6
	ST	1
Diameter of wire (mm)	AL	2,24
	ST	2,24
Diameter of conductor (mm)	Core	2,24
	Overall	6,72
DC resistance (Ω/km)		1,1823

Tab. 1 Parameters of conductor 24-AL1/4-ST1A

As can be seen from the accompanying figure, the conductor temperature was measured in advance, i. before switching on the measuring circuit. The measurement time for unloaded AlFe lane was 10 minutes. The measuring circuit was closed at 10 minutes. The set current value was almost nominal, equal to 125A.

There was a sharp rise in driver temperature between 00:10:00 and 00:20:00. From a conductor temperature of 25-27 ° C at 10 repeat measurements, the temperature increased to 73 to 85 ° C. For averaged driver temperature, this was a temperature rise from 26.2 ° C to 80.5 ° C.

A dampened increase in driver temperature was recorded from 00:20:00. In the case of individual measurements, the temperature rise of the conductor at 00:20:00 for individual measurements ranged from 77 ° C to 85 ° C to 00:50:00 to 77 ° C to 90 ° C. For average temperature values, an increase from 80.5 ° C to 85 ° C was recorded.

FIG. It can be seen that in the case of partial measurements 1-10 the temperature increase was different, but the characteristics of the individual temperature increases were almost identical. This condition was caused by the volatility of the surrounding conditions. In the laboratory, an ambient temperature increase of 1.3 ° C was recorded using a temperature measuring thermometer. Also, the temperature of the driver is influenced by AlFe rope treatment. In the manufacture of these ropes they are treated against oxidation in air to maintain their mechanical properties

To verify the proposed mathematical model for the determination of the static and dynamic part, a calculation was made for determining the temperature of the driver in the dynamic and static parts. From the solution point of view, the determined steady-state conductor temperature was then compared to the average conductor temperature obtained on the basis of the repeat measurements.

By gradual adjustment of the conductor temperature by the iteration method, the determined conductor temperature was TS = 74.6 ° C.

The factors involved in heating and cooling the driver are:

- Heating P_J = 22.88 $W.m^{-1}$
- Warming due to solar radiation PS = 0 $W.m^{-1}$
- Airflow cooling P_C = 19.01 $W.m^{-1}$
- Radiant cooling P_r = 3.88 $W.m^{-1}$
- The difference ΔP is equal to -0.01 $W.m^{-1}$.

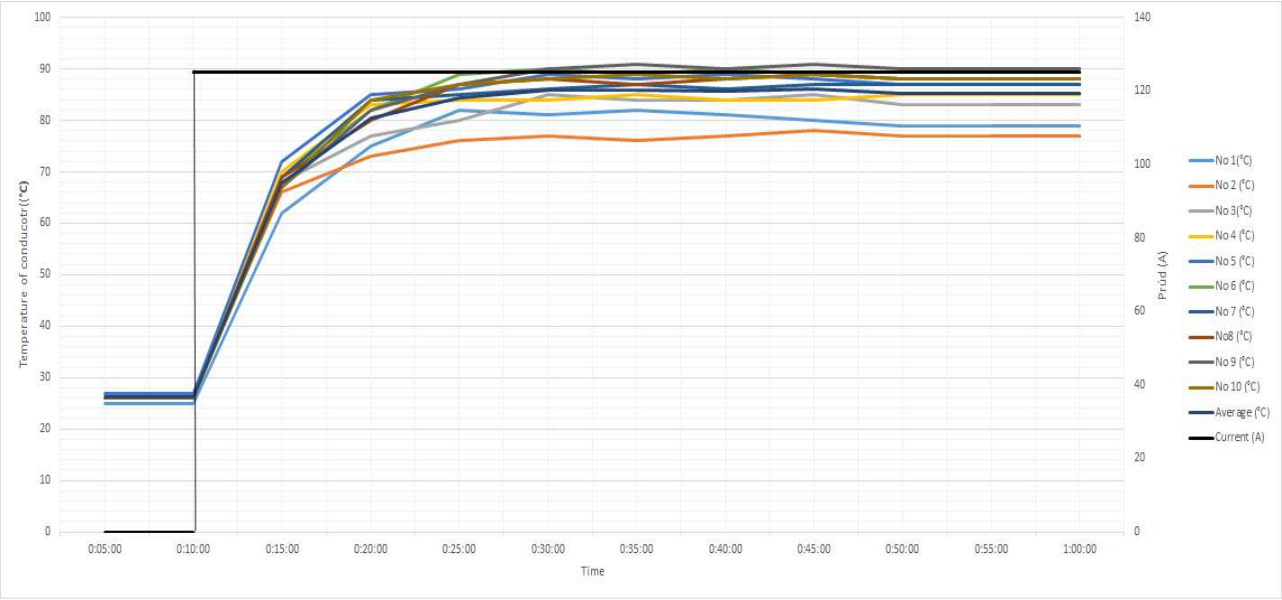


Fig. 3 Conducting the temperature of the 24-AL1 / 4-ST1A wire at a jump load of the rated current

As shown in the enclosed table, when comparing the average conductor temperature obtained from 10 measurements and the calculated conductor temperature, there is a significant difference between these temperatures. The biggest difference was in the first points of measurement, when the difference was up to 67.19%. With increasing measurement time and calculation, the difference between the measured temperature and the calculated temperature decreased to 14.21%. The CIGRE 207 standard defines a temperature difference of up to 20% for smaller wind speeds of 0.5 m / s. Measurement uncertainty also has a significant impact. When comparing the driver's static temperature values, this difference is up to 20% by standard. Due to the low wind speed, the convection is mixed, which is not taken into account by the driver's temperature calculation standard. The graphical representation of the comparison is shown in the following figure.

Average temperature (˚C)	Calculated temperature (˚C)	Difference (˚C)	Difference (%)
26,2	35,5	9,3	26,13
67,9	42,4	25,5	60,04
80,5	48,1	32,4	67,19
84,3	52,9	31,4	59,5
85,8	56,7	29,1	51,27
85,8	59,9	25,9	43,24
85,7	62,5	23,2	37,09
86	64,7	21,3	32,99
85,2	66,4	18,8	28,25
85,2	67,9	17,3	25,51
85,2	69,1	16,1	23,34
85,2	70,1	15,1	21,61
85,2	70,9	14,3	20,22
85,2	71,5	13,7	19,11
85,2	72,1	13,1	18,21
85,2	72,5	12,7	17,47
85,2	72,9	12,3	16,88
85,2	73,2	12	16,4
85,2	73,4	11,8	16
85,2	73,7	11,5	15,68
85,2	73,8	11,4	15,41
85,2	74	11,2	15,2
85,2	74,1	11,1	15,02
85,2	74,2	11	14,88
85,2	74,2	11	14,76
85,2	74,3	10,9	14,66
85,2	74,4	10,8	14,58
85,2	74,4	10,8	14,51
85,2	74,4	10,8	14,46
85,2	74,5	10,7	14,41
85,2	74,5	10,7	14,38
85,2	74,5	10,7	14,35
85,2	74,5	10,7	14,32
85,2	74,6	10,6	14,21

Tab. 2 Comparison of measurement and calculation results for 24-AL1 / 4-ST1A

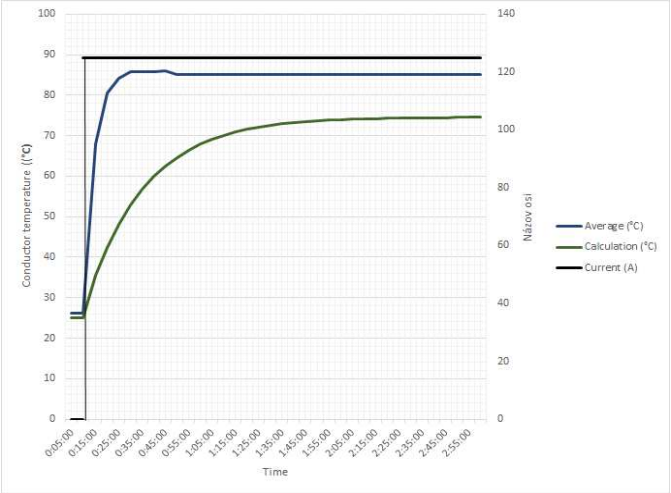


Fig. 4 Graphic comparison comparison for 24-AL1 / 4-ST1A

5. Conclusion

Power lines are one of the most important part of power system. Temperature of the conductor is function of the current value, ambient conditions, type of used conductor and their properties.

Results show that ambient conditions have most influence on the actual value of current capacity. If we can accurately determine the ambient conditions in real time, we can determine the current capacity under these terms and adapt operation of power system or power lines.

Increasing current capacity of existing power lines is one of the way how to operate power system in the short term. On the other side, is necessary to build, expand with new power lines which is however time-consuming and costly.

For the static part, i. steady-state conductor temperature at rated current was found to be 14.21% difference for 24-AL1 / 4-ST1A. In view of such a high difference in measurement and calculation temperatures, the CIGRE 207 standard defines a situation where the air flow rate is less than 0.5m / s difference of up to 20%. In our case, considering the combined measurement uncertainty, this result is excessively accurate. In the dynamic part, i. after switching the measuring circuit, when

the nominal current flowed through the drivers, the difference was significantly higher at 67.19% for the 24-AL1 / 4-ST1A wire. These differences in temperature are beyond the explanation in CIGRE 207. The problem with these parts of the results and its verification is that the CIGRE 207 methodology does not specify a sufficiently accurate calculation option for a state where the wind speed is below m / s in the dynamic part of the load. The missing parameter is the effect of mixed convection.

Acknowledgement

Research described in the paper was supervised by doc. Ing. Ľubomír Beňa Phd. and supported by the Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences by the projects VEGA No. 1/0372/18.

References

[1] M. Kolcun, V. Griger: "Controlling the operation of the power system" (Riadenie prevádzky elektrizačnej sústavy), Mercury - Smékal, Košice, 2003, 288 pages, ISBN 80-89061-57-5.

[2] Böhm V., Popelka A., Vostracký Z.: Ampacity of overhead lines (Ampacita elektrických vedení), CIRED 2010, Tábor

[3] EN 50341-1 ED.2 (333300): Elektrická venkovní vedení s napětím nad AC 1 kV - Část 1: Obecné požadavky - Společné specifikace

[4] Slovenský hydrometeorologický ústav,"Klimatické podmienky na Slovensku" < http://www.shmu.sk/sk/?page=1064>

[5] A. Klenovcanova, T. Brestovic: "Possibilities of utilization of photovoltaic cells for electricity production in Kosice area " (Možnosti využitia fotovoltickych článkov na výrobu elektrickej energie v oblasti Košíc). Acta Mechanica Slovaca 2007, ELFA, s.r.o., 11, 4-D, 511-516, ISSN: 1335-2393

[6] Š. Fecko, J. Žiaran,L. Varga: "Power lines – Overhead power lines" (Elektrické siete - Vonkajšie silové vedenia), SVŠT Bratislava, 1990.

[7] IEEE Standard for Calculation the Current-Temperature Relationship of Bare Overhead Conductors, IEEE Std 738-1993

About Authors

Michal ŠPES was born in 1991. In 2015 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. At present is a Ph.D. student in the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in electric power engineering on subject evaluation of generator exciting outage. His scientific research is mainly focused on research of powerline ampacity system.

Jakub URBANSKÝ was born in 1991. In 2017 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. Currently is a Ph.D. at the same department. His scientific research is mainly focused on research of renewables sources of energy.

Daniel PÁL in 2018 graduated (MSc) at the Department of Electrical Power Engineering on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. Currently is a Ph.D. at the same department. His scientific research is mainly focused on research of optimization of renewable sources .

Michal MÁRTON was born in 1992. In 2016 graduated (MSc) at the Department of Electronics and Multimedia telecommunications. At present is a Ph.D. student in the Department of Electronics and multimedia telecommunications on the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He received a master degree in multimedia telecommunications on measurement with the fibre optic gyroscope system. His scientific research is mainly focused on research of optical communications.

Synthesis and Spark Plasma Sintering of $CaCu_3Ti_4O_{12}$

Jan PETRÁŠEK[1]

[1] Dept. of Electrotechnology, Czech Technical University, Technická 2, 166 27 Prague, Czech Republic

Jan.petrasek@fel.cvut.cz

Abstract. *In this paper we examine dielectric properties of CaCu3Ti4O12 and attempt to discuss the most important conditions affecting its dielectric behavior. At first, we give a quick description of its structure. In the experimental part we describe a not so common synthesis route that showed to give satisfactory results. Finally, we analyze our existing results and compare them with some remarkable achievements found in literature.*

Keywords

Calcium Copper Titanate, Relative permittivity, Loss factor, Dielectric, Capacitors.

1. Introduction

Ceramic capacitors represent an important part of basic electronic components. While these materials show some convenient parameters, scientists still examine new materials, or new combinations of existing ones, in order to improve them. Some examples of such materials can be doped $BaTiO_3$, doped TiO_2, $CaCrTi_2O_6$ or the aforesaid $CaCu_3Ti_4O_{12}$ which show comparatively high values ($\varepsilon_r \approx 10^4$) of relative permittivity.

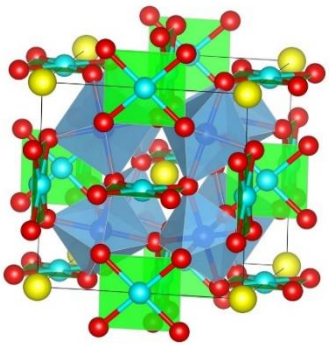


Fig. 1. $CaCu_3Ti_4O_{12}$ structure

One reason why $CaCu_3Ti_4O_{12}$ gained our attention was because titanates were in the past a subject of study at our department. These materials usually belong to the perovskites group, which in general has a formula of ABO_3, being most commonly represented by $BaTiO_3$ and $SrTiO_3$. Analogously the $CaCrTi_2O_6$ and $CaCu_3Ti_4O_{12}$ (Fig.1.) are often labeled as double and quadruple perovskites, respectively. (There are also a lot of members of the triple perovskites group, but up to now it seems that none of them has any competitive parameters.)

2. Experimental

For the experimental part, the not so common way of synthesis in a molten salt was chosen. $CaCO_3$ (99 %, Sojuzchimexport), CuO (99%, Alfa Aesar) and TiO_2 (99%, Alfa Aesar) were bought in powder form. These powders were then weighted according to the stoichiometry, mixed together and blended in ethanol by a hand stirrer for approximately 15 minutes. The ethanol was let to evaporate and these powders mixed with sodium chloride (laboratory grade, Verkon). Together they were put into a refractory crucible and heated in a resistance furnace for 2 hours with temperature set to 750°C. After cooling down, the product was dissolved in warm distilled water and sodium chloride solution separated with filtration paper. This process was repeated 5-10 times, after which the water electrical conductivity raised up to 4 orders of magnitude.

The resultant powder was given to the Institute of Plasma Physics ASCR for X-ray diffraction analysis and Spark Plasma Sintering. In total 8 samples were sintered under different conditions and their dielectric parameters measured (Fluke PM6306, Isotech Venus 4852) at FEE CTU.

3. Results

After the dielectric measurements, our samples were examined with scanning electron microscope in order to find correlations with microstructure. One example of material imperfection observed in the early sintered samples are some very heterogeneous areas (Fig. 2). Here on a grey background rich in titanium we can find some dark pores and light spots representing areas rich in copper. In agreement with this fact, the average relative permittivity of this sample was around 160, with the dielectric loss factor being as high as approximately 1.6.

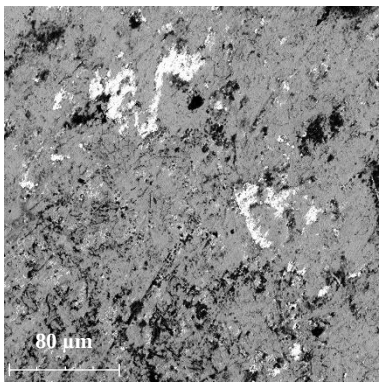


Fig. 2. Microstructure of sintered $CaCu_3Ti_4O_{12}$

Annealing in various gases is typical procedure for improving the material's structure. In this case, however, we observed some ambivalent results. On one side, the microstructure has changed; finer grains are now visible. These grains look similar to those observed in [1], which we can regard as a promising result (see Fig. 3). On the other side, this change was also accompanied with pores coalescence.

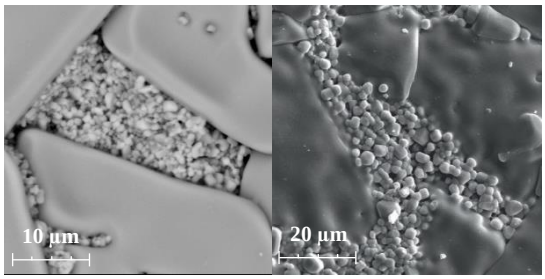


Fig. 3. Comparison of our result (left) with [1]

While we achieved a relative permittivity comparable with our assumptions, we can see that it is strongly frequency dependent (Fig. 4.). It also a challenge for the future to improve the loss factor values which now do not descent below 0.2.

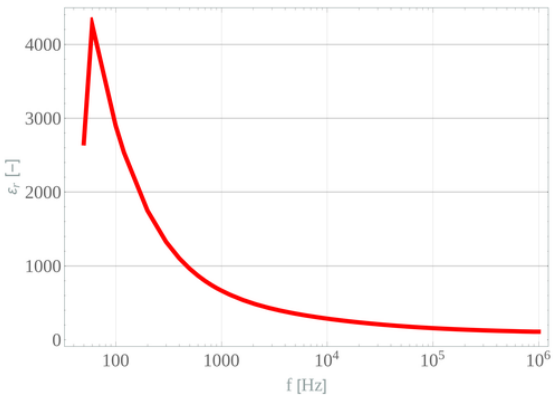


Fig. 4. Frequency dependence of relative permittivity

4. Conclusion

Dielectric parameters of $CaCu_3Ti_4O_{12}$, their relation to the materials' microstructure and optimal setting of Spark Plasma Sintering were studied. The preliminary results show that sintering temperature below 1000°C should be set at SPS process. Longer dwell time at the maximum temperature is more efficient for receiving defect-free compacts than high temperature. The role of pressure and its application schedule is not clear enough at this time. Contamination of the dielectric material by carbon coming from the apparatus was successfully avoided.

Acknowledgements

This research was supported by the Grant Agency of the Czech Technical University in Prague, No. SGS18/070/OHK3/1T/13.

The author would also like to thank to Pavel Ctibor and Josef Sedláček for their help and supervision.

References

[1] PUCHMARK, Chompoonuch. Improvement in dielectric and mechanical performance of CaCu3.1Ti4O12.1 by addition of Al2O3 nanoparticles. *Nanoscale Research Letters* [online]. 2012, **7**(1), 68-68 [cit. 2019-03-24]. DOI: 10.1186/1556-276X-7-68. ISSN 1931-7573.

[2] AHMADIPOUR, Mohsen. A Short Review on Copper Calcium Titanate (CCTO) Electroceramic: Synthesis, Dielectric Properties, Film Deposition, and Sensing Application. *Nano-Micro Letters* [online]. 2016, **8**(4), 291-311 [cit. 2019-03-24]. DOI: 10.1007/s40820-016-0089-1. ISSN 2150-5551

[3] LIN, He. Tuning the nonlinear current-voltage behavior of CaCu3Ti4O12 ceramics by spark plasma sintering. *Ceramics international* [online]. 2018, **44**(7), 8650-8655 [cit. 2019-03-24]. DOI: 10.1016/j.ceramint.2018.02.089. ISSN 0272-8842.About Authors

About Author

Jan PETRÁŠEK born in 1989, Prague, has graduated from Czech Technical University in the field of Power Electrical Engineering and currently is a postgraduate student at the Department of Electrotechnology. In his thesis he examines structures and dielectric properties of perovskite related materials.

An introduction to voltage gain of LLC resonant converter

Ing. Filipp Frolov

[1] Dept. of Electric Drives and Traction, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic

frolofil@fel.cvut.cz

Abstract – *the aim of this paper is to present the model of LLC resonant converter, this model is based on quantitative analysis of the steady-state operation of the circuit. This analysis follows the first harmonic approximation (FHA), which tremendously simplifies the model, leading to a linear circuit, which can be dealt with through the classical complex ac-circuit analysis.*

One of the major benefits of the LLC resonant topology is the ability of the Power transistors to be soft-switched. Soft-switching greatly reduces switching losses, allowing for LLC use in high-frequency applications. Operating at higher frequencies reduces the size of passive component, such as transformer and inductor, considerably and permits higher power densities.

Keywords

LLC resonant converter, first harmonic approximation, Ac-circuit analysis, transfer function, full-bridge.

1. Introduction

The LLC resonant converter is widely used in medium-to-high-power conversion applications because of its high efficiency, low level EMI emissions and its ability to achieve high power density. The purpose of this paper is to provide a quantitative analysis of the steady-state operation of the topology that can be easily translated into a design procedure.

2. FHA Circuit Model

The circuit of LLC resonant converter is shown in Fig.1. The circuit consists of three parts. Square-wave generator - four power transistors (S1-S4) are configurated to produce a bipolar square-wave voltage. Resonant Tank - the resonant network is formed by capacitor C_r, inductors L_r and L_m. In particular, L_m represents the transformer's magnetizing inductance. The third part is rectifier and filter – on the secondary side of a converter, the rectifier consists of four diodes for full-wave rectification and an output capacitor C_f to smooth the rectified voltage to the load R_L.

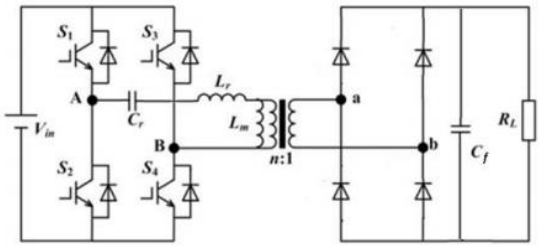


Fig. 1. Circuit of Full-bridge LLC resonant converter

Fundamental harmonic analysis represents a consolidated method to derive the transfer function of the resonant tank and the output rectifier stage. The FHA frequency domain method sacrifices model accuracy to simplify the topology through sinusoidal waveform approximations and AC equivalent circuits. Using the Fourier series, we can break up the square-wave input voltage and output voltage through their harmonic components. Fig.2. is described the equivalent circuit of the LLC resonant converter. The equations 1. and 2. are describe the substitutions.

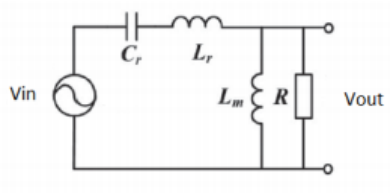


Fig. 2. AC equivalent circuit for the LLC resonant converter

$$R = n^2 \frac{8}{\pi^2} R_L \tag{1}$$

$$n = \frac{N_p}{N_s} \tag{2}$$

3. Voltage gain of LLC circuit

From the equivalent AC circuit in Fig.2. we can define the following parameters (2), (3). these parameters simplify the circuit analysis.

$$Z_1 = j\omega L_r + \frac{1}{j\omega C_r} \tag{3}$$

$$Z_2 = (R \parallel j\omega L_m) = \frac{j\omega L_m R}{j\omega L_m + R} \tag{4}$$

Define the voltage gain (5) of LLC resonant converter as transfer function from input voltage to output voltage.

$$G\left(\frac{f_s}{f_r}, Q, k\right) = \frac{V_{in}}{V_{out}} = \frac{Z_2}{Z_1 + Z_2} \tag{5}$$

Simplify the (5) and define the next parameters: Q – quality factor on the load, k – inductive ration

$$Q = \frac{\sqrt{L_r}}{R\sqrt{C_r}} \tag{6}$$

$$k = \frac{L_m}{L_r} \tag{7}$$

The voltage gain function is describe in (8)

$$G\left(\frac{f_s}{f_r}, Q, k\right) = \frac{1}{\sqrt{\left(1 + \frac{1}{k}\left(1 - \frac{{f_r}^2}{{f_s}^2}\right)\right)^2 + Q^2\left(\frac{f_s}{f_r} - \frac{f_r}{f_s}\right)^2}} \tag{8}$$

f_s - switching frequency

f_r – circuit resonant frequency

Fig. 3. shows a family of plots of voltage gain versus normalized frequency. For different Q values at the inductive ration k = 20.

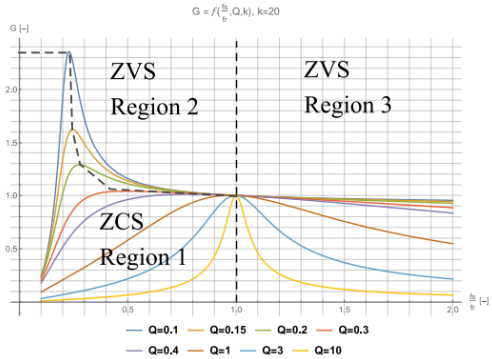


Fig. 3. Operating regions for an LLC resonant converter

References

[1] FROLOV F. , master's thesis – Design of Resonant Power Converter with Wide Range of Supply Voltage, K13114 - 2018

[2] F. Canales, P. Barbosa, and F.C. Lee: A wide input voltage and load output variations fixed-frequency ZVS DC/DC LLC resonant converter for high-power application, 2002 37th IAS Annual Meeting of Industry

[3] B. Yang, F. C. Lee, A. J. Zhang, and G. Huang. "LLC resonant converter for front end DC/DC conversion." In Applied Power Electronics Conference and Exposition, 2002. APEC 2002

[4] B. Lu, W. Liu, Y. Liang, F. C. Lee, and J. D. Van Wyk. "Optimal design methodology for LLC resonant converter." In Applied Power Electronics Conference and Exposition, 2006. APEC'06

[5] Steady-state Analysis of the LLC Resonant Converter, Applied Power Electronics Conference and Exposition, 2001. APEC 2001. Pages: 728 - 735

[6] S. De Simone, C. Adragna, C. Spini and G. Gattavari - Design oriented steady state analysis of LLC resonant converters based on FHA

[7] Xinbo Ruan and Jiangang Wang - Calculation of the Resonant Capacitor of the Improved Current-Doubler-Rectifier ZVS PWM Full-Bridge Converter - IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, VOL. 51, NO. 2, APRIL 2004

About Author

Ing. Filipp Frolov was born in Perm, Russian Federation. He was graduated from Czech Technical University, Prague, 2018. He is currently working toward the Ph.D. degree in power electronics with Czech Technical University in Prague.